technopolis |group|

November 2010

REF Research Impact Pilot Exercise Lessons-Learned Project: Feedback on Pilot Submissions

Final report

Table of Contents

Table of Figures

REF Research Impact Pilot Exercise

Lessons-Learned Project: Feedback on Pilot Submissions

Draft final report

1. Summary

1.1 This report

This report to HEFCE presents feedback from the 29 pilot universities and colleges (HEIs) that participated in the REF Research Impact pilot exercise, bringing together the insight and lessons learned from the submission process. Feedback was gathered through interviews and written contributions, ahead of the Assessment Panels' evaluations of those submissions.

This report is concerned with the specific experiences of the Pilot HEIs' and the lessons that can be learned. It will complement the feedback being compiled by the chairs of the five Assessment Panels. It will contribute to helping HEFCE define a robust and workable method for assessing impacts, through the forthcoming REF, and will provide an additional reference to HEIs when making their preparations for the REF. The report is structured around the questions posed to the Pilot Institutions and the chronology of the submission process and their experiences, beginning with a discussion of the appropriateness of the HEFCE submission guidelines and briefings.

1.2 HEFCE Guidance on submissions

The HEFCE Guidance and briefings were well received by institutions, in the main. While the HEFCE Guidance proved helpful overall, most Pilot Institutions expressed some minor concerns. Most commonly, people criticised four headings within the impact statement, which were confusing and seemingly repetitious. Several key concepts were ambiguous, 'interim impact' as compared with 'final impact', for example, or 'reach' and 'significance'. Lastly, most Pilot Institutions that made an English submission noted that the 'common menu' of indicators in the appendices to the Guidance was light on measures relevant to the social sciences and the humanities.

For the REF proper, the Pilot Institutions asked that HEFCE make full use of this experience to expand its instructions, minimise ambiguity and bring the material to life with worked examples.

1.3 The submission process

HEFCE chose not to prescribe the process by which submissions should be prepared. However, a standard approach did emerge, with central administrators project managing the exercise, senior departmental academics taking the lead in drafting impact statements and case studies, and high-level institutional committees peer reviewing material and providing tactical advice.

There were variations in the way things were organised across institutions and units of assessment, with differences evident on two dimensions:

- The division of labour between the central team of (research) administrators and the academics, with a significant minority of all institutions looking to their central administrative teams to play a prominent role, possibly preparing the first drafts of impact statements or case studies, with senior faculty and principal investigators quality assuring the work
- The extent of involvement of all members of the faculty. In several Pilot Institutions, one or two senior academics volunteered to carry out most of the

work. Elsewhere, Pilot Institutions adopted a more bottom-up approach involving the whole department

Pilot Institutions were equivocal about the best approach. However, people were settling to the idea that there needed to be a strong and substantial contribution by senior research administrators, and their support staff, to minimise the burden on key academic staff, and a faculty-wide input to the long-listing and selection of impacts and a more cooperative approach to drafting.

There were few evident strong differences that were subject-specific.

Almost half the Pilot Institutions took the opportunity to run a wider programme of internal events and workshops to raise awareness about research impact and a significant minority ran shadow impact assessment processes for other subject areas.

Each of the 29 Pilot Institutions was able to prepare a compliant submission or submissions, with no expressed concerns about the completeness or quality of the material provided, which suggests a generic model is a practicable solution for the wider HE community.

1.4 Challenges overcome

The most widely reported challenge was conveying HEFCE's meaning of research impact to their respective academic groups and, in particular, that socio-economic impact was a much broader concept than economic impact.

It proved to be non-trivial for coordinators to explain HEFCE's view of non-academic impacts to their academic audience, and substantial effort was devoted to general communication as well as to bilateral exchanges around specific drafts of impact statements or impact case studies.

Most Pilot HEIs responded by developing additional illustrative case material and grasping opportunities to involve colleagues in constructing self-evidently good and relevant examples of non-academic impacts.

The biggest challenge was the need to acquire evidence of the reach and significance of a given impact. There were many practical issues that stood in the way, and the response of the majority, for the pilot exercise at least, was to firstly focus on the more obvious cases and, secondly, to use whatever narratives, references and statistics that came to hand readily.

Respondents expect the evidentiary challenge to become less problematic in the medium term, as academics acclimatise to these requirements and begin to log evidence routinely. However, several contributors argued that while new behaviours and systems will improve the evidence base, and the ease with which it can be assembled, intrinsic difficulties will remain, by virtue of the manifold, diffuse and long-run nature of research impact. In light of these observations, we expect many institutions will still select their portfolio of REF case studies in part based on the availability of evidence.

1.5 Quality of the submissions

The great majority of Pilot Institutions were satisfied with the quality of the impact accounts they had prepared. There was no general sense of the process or templates having hamstrung institutions, albeit this judgement might need to be revisited following publication of the results.

There were challenges to overcome, to do with uncertainties over the information requirements and the timeframe within which submissions had to be made ready. One might expect the impact statements and impact case studies to be a little stronger on average in the REF proper.

The impact statement split opinion, with most suggesting it was time poorly spent and a distraction from the important case studies. By contrast, a minority saw the impact

statement as a platform for the wider involvement of faculties and a systematic starting point for selecting case studies. Several Pilot Institutions stated that the faculty-wide review had already been of strategic value, outside the pilot exercise.

The case study approach, by contrast, was widely endorsed as being the most appropriate methodology with which to illustrate research impacts and research-impact pathways.

1.6 Lessons learned

All of the 29 Pilot Institutions reported that they had learned valuable lessons as a result of participating in the pilot exercise. A small minority qualified their remarks, adding that the most important insights would follow feedback on their submission.

Most were strongly positive about the experience and collectively cited some half-dozen institutional benefits, wherein institutions had:

- A much better grasp of what is meant by non-academic research impact, in particular its broad scope
- A much better sense of what will be required in order to detail the nature and extent of research impacts, and how to avoid at least some of the potential pitfalls in preparing submissions
- A new perspective on many excellent things that have happened as a result of research undertaken within the university

Most Pilot Institutions reported good progress in respect to alerting their academic colleagues as to the importance of more general reflection and recording of the outcomes and impacts of their research, and a significant minority suggested the exercise had demystified the whole concept and won over at least some sceptics.

The great majority of Pilot Institutions concluded that the inclusion of a research impact element within REF will require additional effort on their part to develop departmental and institutional procedures and information systems, better tailored to recording impact-related data and material.

Considering this feedback, it seems likely to the authors that the first developments will be procedural, and that most Pilot Institutions will:

- Ask all departments to develop impact case studies, to expand the total institutional catalogue of reference cases and to help build institutional capacity
- Ask all researchers to think more about what good things happen outside academia as a result of their work, inviting people to record any relevant material or data as and when it arises

The great majority stated that their institution was broadly content with the approach adopted in the pilot exercise, and would not have wanted to do things differently, even with the benefit of hindsight. This suggests that the process description captured through this feedback exercise can be used by HEFCE as a model process for the HE community overall.

A small minority stated they would do things a little differently, were they to do the exercise again. It seems that resourcing was an issue in several Pilot Institutions, with a sense that they had relied too heavily on a small number of individuals and that a small central team – with significant support from senior research administrators – would have reduced stresses and pinch points amongst the costly, busy senior academic staff.

1.7 Cost to prepare a submission

Twenty-five of the 29 Pilot Institutions supplied us with data on the number of staff-days their submissions took to prepare.

At the most general level, it took an average of 57 staff-days for an institution to coordinate and prepare a full submission. This figure masks important structural differences, reflected in the range of estimates for total staff-days, with the lowest estimate being 11 staff-days (covering 27 FTEs and involving four case studies) and the highest being 194 staff-days (covering 519 FTEs and 53 case studies).

Adjusting these aggregate data with data on the numbers of researchers submitted (in full-time equivalents or FTEs) produces an estimate for the average amount of effort involved in preparing a submission, which is 0.5 staff-days *per FTE*.

Within this overall cost, the average cost of preparing case studies was 0.3 staff-days per FTE. The average cost of preparing the overviews and impact statements was 0.1 staff-days per FTE. The balance of costs, 0.1 staff-day per FTE, was given over to the time taken to interact with HEFCE, brief staff and coordinate the submission.

Those preparing submissions involving large numbers of case studies expended a lower amount of effort *per case study* than those that only produced a small number. The average amount of effort per case study for those producing ten or fewer cases was 5.9 days, while the average effort per case study for those institutions producing more than ten case studies was 2.7 days. It is therefore likely that within the context of a full REF exercise, the effort involved in preparing each case study would be toward the lower end of the estimates provided here.

The smallest amount of effort per impact statement for any given institution was one staff-day and the largest was 23 staff-days.

The composition of administrative and academic time varied from pilot to pilot, with an even split, 50:50, at the lower end, through to an 80:20 split in favour of academic contributions, at the upper end. The latter split was more likely amongst those Pilot Institutions making smaller submissions. Institutions making larger submissions had a proportionally bigger input from central research and planning services, perhaps 70:30 in favour of academic contributions.

Overall, the feedback makes clear that a REF Impact Assessment module will impose real additional costs on institutions in respect of their establishment and ongoing operation of new procedures and systems and the preparation of impact accounts to accompany their excellence and environment accounts. However, none of the Pilot Institutions considered the submission costs to have been unreasonable and a significant minority expect those additional costs to be offset by institutional benefits arising from changed behaviours.

The principal exception as regards value for money was the impact statement, where a significant minority suggested that the document had been costly to produce and was of uncertain value to the institution or the assessment process.

1.8 Subject-specific issues

Looking across the feedback from 29 institutions, HEFCE's generic approach to the submission process has worked reasonably well, with no insurmountable difficulties experienced in any subject.

There were subject-specific challenges. These concerned English most often, where impacts tended to be more conceptual than instrumental, so intrinsically difficult to convey and dimension. The response was twofold, to focus on the small number of *obvious* impacts and otherwise to develop case studies of interim impacts (notable achievements, that might be expected to contribute to a wider societal impact in the fullness of time), which used narrative accounts, qualitative in the main, and relied on self-evidently noteworthy events as a proxy for future impact.

Not all English departments struggled with the exercise, and there was no clear split between humanities and the other four subject areas in terms of the degree of difficulty encountered. Indeed, in several cases, Pilot Institutions reported that their 'non-English' unit had struggled with the exercise more than their English

department. There were also widely reported challenges in each of the other four subject areas, with a significant minority of contributors suggesting the exercise had ultimately focused on impacts linked with their more applied research while their theoreticians and basic researchers had found it harder to identify and dimension specific socio-economic impacts attributable in some critical manner to their own work.

Looking across the 29 accounts, it appears that each of the five subject areas might have what amounts to an impact *signature*, a characteristic profile of impact pathways and types of impacts, particular to the subject. For example:

- Clinical Medicine and commercial confidentiality
- Earth Systems and Environmental Science and the importance of interim impacts
- English and the centrality of individual scholarship conducted over a lifetime
- Social Work and Social Policy and the valuation of the negation of existing policies
- Physics and the attribution challenges posed by working on international projects

The Pilot Institutions were broadly content that the use of subject-specific Assessment Panels would permit the REF to cope with these kinds of disciplinary particularities, such that judgements and grades can be appropriately normalised.

1.9 Behavioural implications

The very great majority of Pilot Institutions expect the introduction of the REF impact assessment component to bring some change in behaviour and, looking across the feedback, one can see anticipated changes at a number of levels: individual researchers, institutions and possibly disciplines.

At this stage, however, few of the Pilot Institutions expect the behavioural changes to be as profound as those brought about by the Research Assessment Exercise, when it was first introduced.

A significant minority was vocal in its support for the introduction of a research impact dimension to REF, expecting this to focus the minds of all staff on doing good, relevant work, as well as perhaps helping to boost the 'within-institution' status of researchers and research groups with a more applied focus. This group of Pilot HEIs believe impact assessment will produce positive feedback, strengthening UK research quality *and* utility over time.

However, there were several detractors; people who believe a focus on non-academic impacts might change the composition of UK research and, in particular, bring about a diminution in the total amount of excellent, fundamental research being performed. For this group, the reputation of UK science could be harmed as a result of such a changed composition and, perversely, one should also expect a reduction in the total economic benefits attributable to UK research.

There was a sense, more widely expressed, that we ought to be cautious as regards the weighting of elements within the REF, given the newness of the impact assessment approach. No one offered a suggestion as to what the share ought to be, although several acknowledged the challenge HEFCE will face in determining the appropriate weighting

1.10 Advice to non-pilot HEIs

The most widespread advice to non-pilot HEIs is that they should begin their preparations for the REF Impact Assessment immediately.

The pilot exercise confirmed that the act of gathering evidence on research impact is a largely new endeavour, with little institutional infrastructure available to support the process and heavy reliance on the personal knowledge of senior academics.

This embodied quality-led people to suggest that HEIs should do more to write down what is already known about past impacts and, going forward, to explore cost-effective options to encourage collection of impact-related material and evidence.

The Pilot Institutions foresee the need for a major communications exercise too, with each university needing to launch conversations with all faculties in all units of assessment. Equally, the Pilot Institutions anticipate this communication exercise requiring rather more than a communiqué and an accompanying event, given that many academics remain sceptical and will view the REF proposals with a mixture of anxiety and antagonism.

1.11 A final comment

Overall, the Pilot Institutions have come away from the exercise feeling much more comfortable with the notion of research impact and its relevance to their institution, as well as increasing confidence in their ability to document research impact.

People commented favourably on HEFCE's openness in running the exercise and complimented the REF Team on its willingness to debate issues. Most respondents also praised their fellow Pilot Institutions for their candour in discussing their experiences and readiness to offer advice to others.

Institutions had also come to a new understanding of their many and various specific contributions to the wider world, and many were pleasantly surprised at the results.

1.12 Conclusions and recommendations

In conclusion, the feedback from Pilot Institutions has confirmed the feasibility of the approach tested through the pilot exercise; it is clear that HEIs can document non-academic impacts and that in doing so a great majority will derive insight and local benefits. The use of subject-specific Assessment Panels means a generic model should work for all disciplines.

The feedback alone is not sufficient to determine value for money. However, it does suggest the model trialled here is proportionate, given the scale of the research endeavour. Moreover, it is likely to produce increased benefits to institutions, and possibly the UK overall, through positive feedback, which should offset and possibly greatly exceed any additional costs to the national research system. There is a risk of negative feedback, however, which would effectively increase the costs of this form of scrutiny and reduce the attractiveness of the impact assessment proposal.

We offer a number of recommendations to HEFCE for it to consider when synthesising the lessons learned through the submission process with those derived from the assessment process, and which might help strengthen the REF proper.

Overall, we recommend HEFCE move forward with the generic model tested through the pilot, retaining the impact case studies as the centrepiece of each submission, for assessment by subject-specific Assessment Panels.

We recommend HEFCE consider its options for retaining the impact statement somewhere within the REF process.

HEFCE might consider the possibility of switching aspects of the impact statement to the Research Environment component of the REF, rather than eliminating it altogether. While the majority of Pilot Institutions found the overview largely duplicative and unhelpful to their preparations, this was not always the case. The feedback has led us to believe the impact statements should add value, even if it is not scored directly, by more systematically revealing the spectrum of impact types, cataloguing a majority of all impacts (inventory) and providing an analytical framework for selecting the best mix of high-impact cases.

On the HEFCE Guidance, the feedback points to several areas where further work by HEFCE might be beneficial to any future guidelines:

- Adding a glossary of terms
- Redrafting the text in several of the key sections to improve legibility and minimise the risk of misinterpretation
- Explaining ambition levels, in terms of the scope of a case study and the weight of evidence expected, as these two factors will determine costs to a large extent
- Using the pilot impact statements and case studies to develop illustrative material
- Using the Assessment Panel's feedback and pilot submissions to extend the draft common menu of indicators

In terms of the REF proper, we recommend HEFCE:

- Implementing additional support measures, such as a communications pack, to facilitate institutions' internal communications
- Defining a timeframe for the submission process that allows sufficient time for institutions to consult their academics and engage research users more fully

The preceding recommendations are directed to HEFCE. In the main, the feedback confirmed that individual Pilot Institutions have derived much insight and learning from the exercise, which might be of significance to the wider community.

2. Introduction

2.1 This report

This report presents feedback from the 29 Pilot Institutions that participated in the REF Research Impact pilot exercise, regarding the submission process.

This report is concerned with the specific experiences of the institutions and the lessons that can be taken away from them and will complement the feedback being compiled by the chairs of the five Assessment Panels. It will contribute to helping HEFCE define a robust and workable method for assessing impacts, through the forthcoming REF, and it will provide a reference to help guide any HEIs involved in making a preparation for the REF.

2.2 Background

HEFCE launched a pilot exercise in autumn 2009 to trial an impact assessment methodology, which is intended to inform the operational design of an impact-assessment procedure for the proposed Research Excellence Framework (REF) covering the entire UK higher education sector. The assessment of research impact will be one of three distinct elements of the REF, being judged alongside research excellence and research environment, contributing 25% towards the overall outcome (as compared with 60% and 15% for quality and environment).

The focus is on socio-economic impacts of any type and in any realm, realised outside the academic community. The impacts should have been evident at some point during, or indeed throughout, a given time window (2005 – 2009) and should relate unambiguously and decisively to one or other aspects of a university department's research. This high-quality underpinning research must have produced a formal output at some point in the period since 1993.

The pilot exercise was carried out to test this approach through the combined efforts of 29 higher education institutions (HEIs), submitting evidence of impact (an impact statement exemplified through impact case studies) to be assessed and graded by expert panels in the following five (REF) Units of Assessment (UOAs):

- Clinical Medicine
- Physics
- Earth Systems and Environmental Sciences
- Social Work and Social Policy
- English Language and Literature

2.3 The feedback process

The Pilot Institutions were invited to answer ten broad questions, which have been used to structure the analysis and presentation of the findings in this report.

The questions were developed by Technopolis in conjunction with the HEFCE REF team and discussed with the Pilot Institutions at a briefing event in London. They were trialled through preliminary interviews with the lead administrators at the universities of Cambridge and Exeter.

Our methodology was designed to combine an interactive approach of telephone interviews with a majority electing to provide a written response in order to facilitate input by academic groups and central services in a single submission. All 29 Pilot Institutions made time to contribute to the feedback exercise. We also made contact with a number of departments and had discipline-level exchanges. The analysis was therefore carried out at two levels, both institutional and subject.

The feedback was primarily of a qualitative nature, concerning the individual experiences of the departments and institutions in the trial process. However, we also sought to combine this with quantitative data, particularly on the timing and cost of preparing the submissions and permitting some simple descriptive statistics. The questions are reproduced in the appendices to this paper.

For reference, Figure 1 tabulates the 29 Pilot Institutions against the numbers of impact case studies that they submitted, by subject. There were 26 dual submissions and three single submissions, or 55 units in total, split reasonably evenly across the five subject areas. Together, the Pilot Institutions prepared and submitted 324 case studies, again with a good spread across disciplines, albeit the 40 case studies prepared by Imperial College's medical school did somewhat skew the subject mix. The number of case studies varied greatly amongst the Pilot Institutions with 16 submitting fewer than ten cases and four submitting more than 20; Imperial College stands apart from the rest, with its total count of 53 cases, and is worthy of a special mention for this very substantial contribution to the process.

Figure 1Numbers of impact case studies submitted, by Pilot Institution and subject

	Clinical Medicine	Physics	Earth Systems	Social Work	English	UOAs Submitted	Cases Submitted
University of Bristol	7			5		2	12
Brunel University			2		2	2	4
University of Cambridge		15			8	2	23
Cardiff University	13				5	2	18
De Montfort University				3	2	2	5
University of Dundee	9				2	2	11
University of Durham		7	4			2	11
University of East Anglia			8	2		2	10
University of Exeter	4*				4	2	8
University of Glasgow	17		4			2	21
Imperial College	40	13				2	53
Keele University		1		4		2	5
Kingston University					2	1	2
Lancaster University		3			6	2	9
University of Leeds			8	4		2	12
Liverpool John Moores University		2			2	2	4
University of St Andrews		4			3	2	7
London School of Economics				5		1	5
London South Bank University				2		1	2
University of Stirling			2	3		2	5
Swansea University		3		2		2	5
University of Manchester			4		4	2	8
University of Oxford	24		4			2	28
University of Plymouth	4*		4			2	8
Queen Mary University London	10				4	2	14
University College London		11			3	2	14
University of Ulster			3	2		2	5
University of Warwick	4	6				2	10
University of York		3		6		2	9
Total no. of units making a submission, by subject	10	11	10	11	13	55	
Share of all units submitting, by subject	18%	20%	18%	20%	24%		
Total number of case studies, by subject	128	68	43	38	47		324
Share of all case studies, by subject	40%	21%	13%	12%	15%		

Note: * Plymouth and Exeter made a joint submission for Clinical Medicine

3. Guidance on submissions

3.1 Background

The first question addressed in the feedback related to the appropriateness and clarity of the written guidance issued by HEFCE in a published paper entitled, REF Impact Pilot Exercise: Guidance on Submissions (19 November 2009), which encompassed

- General background and aims of the pilot
- The information to be submitted by the Pilot Institutions
- An elaboration of the scope of the submission, with definitions and threshold criteria
- An explanation of the assessment process and feedback

The document was to be read in conjunction with the relevant sections of the consultation report (HEFCE 2009/38).

The written guidelines were supported through a series of briefing events involving representatives from all of the Pilot Institutions, which served both as an opportunity for HEFCE to explain its purpose and procedures and also a form of peer-learning activity for the Pilot Institutions.

3.2 Overall

The great majority of Pilot Institutions considered the HEFCE Guidance to have been reasonably helpful and a significant minority complimented HEFCE on what they judged to be a good first edition of instructions for what is an entirely new and challenging aspect of institutional performance measurement in the HE sector. In addition, there were positive remarks made about the briefing events and in particular the ad hoc advice from the HEFCE REF team.

However, the feedback was not entirely positive. Several institutions considered the guidance to have been insufficient, even for a pilot exercise, with what one respondent referred to as 'skeletal instructions expressed in a vague and ambiguous fashion'.

3.3 The submissions

3.3.1 Overview information and impact statement

Pilot Institutions tended to provide feedback on the guidance relating to each of the principal elements in their submission, which is to say the overview, impact statement and impact case studies.

Most had little to say about the overview. However, where respondents did comment it was almost exclusively negative. This challenge was driven by an expectation that significant parts of the overview would simply duplicate components of the information required for the research environment component of the REF2014 submission and, as such, were deemed to be unnecessary to this trial. Indeed, HEFCE uploaded institutions' RA5 documents, submitted to the 2008 RAE, to facilitate the re-use of existing research-environment statements in order to streamline the preparation of an up-to-date overview of a units' research in the period under review (the time window was different for the RAE).

While this was a largely duplicative exercise, discussions of costs suggest that it required very limited additional effort and, as such, HEFCE will want to reflect on the value derived by the Assessment Panels: does a bespoke presentation of information

about the research environment add value and is it worth the additional effort to prepare a separate presentation of these common data?

In practice, a majority of respondents addressed the bulk of their feedback to the impact statement, that part of the overview information will be entirely new for REF. Two concerns emerged, both relating to the clarity and legibility of the guidelines:

- The HEFCE Guidance left Pilot Institutions feeling rather uncertain as to the role of the impact statement within the overall assessment process. Was it going to be judged in a formal sense, for example, or simply read as an introductory text to familiarise assessors with the unit in question?
- The questions in the impact-statement template were considered to be poorly drafted.[1] Feedback shows several instances where Pilot Institutions realised – through offline discussions at briefing meetings – that they had arrived at contrary understandings of the same question. The questions in sections 1 and 2 were felt to have asked for essentially the same information, while the question in section 4 proved confusing on at least two counts (e.g. what kinds of activities were foreseen that were materially different to the 'opportunities' or 'types of interactions' referred to in Q1 and Q2; and were university-level initiatives eligible or only activities that were specific to the given unit?). Section 4 was also seen as duplicating information that was expected to be provided in the Environment section of a future REF submission.

A significant proportion of Pilot Institutions took the view that the impact statement was a challenging and costly account to prepare and that such effort was unwarranted if the document was to be used simply as a scene-setting piece for the Assessment Panel. This sub-set of institutions suggested the costs outweighed the benefits, and that the impact assessment had produced little insight of value to the unit or institution.

In marked contrast, a significant minority of Pilot Institutions took the opposite view, noting the importance of the impact statement. It was considered to be especially important to the smaller units, where the allotted number of impact case studies was deemed to be insufficient to reveal the range of contributions being made. In several cases, the impact statements were reported to have been a source of insight, revealing the breadth of their unit's wider contributions, really for the first time, as well as facilitating the process of identifying a long list of case study candidates from which to choose the best portfolio. However, even here, people asked for clarification over the role of this overarching statement and several respondents recommended that, given the effort involved in preparing an impact statement, it should be subject to a formal assessment with the results feeding into the overall impact profile for the unit.

No one went as far as to suggest how an impact statement might be scored such that it might be added to the other scores and thereby factored into the impact profile. Given it is an overview, it is not clear whether one could use the same assessment criteria as for the case studies, reach and significance. Arguably, the assessment of the full account – rather than selected cases – would need to weigh the number and range of types of impact (interim and final), albeit adjusted for the size of the unit and normalised against some kind of subject-specific reference.

[1] For ease of reference, the four questions and prompts were: 1. During the assessment period, what range of opportunities have been identified to apply, exploit or build on the unit's research findings? 2. Describe the range and types of interactions with research users (or potential users) during the assessment period. 3. Outline the full range and significance of impacts or benefits occurring during the assessment period that the unit's research activity has contributed to. 4. What activities are undertaken currently within the unit to build on research to secure future impacts or benefits?

3.3.2 Impact case studies

The great majority of respondents expressed strong support for a case study approach to illustrating research impact. Moreover, people liked the simple, three-point structure of the case study template and the subsidiary points under each heading.

A minority asked for the order of the template headings to be reversed, to begin with a description of the research before describing the resultant impact, which reflects the chronology of events. This suggestion seems to follow the experience of how best to engage researchers in the process, many of whom are reported to have found it very difficult to begin with an impact and in several cases coordinators report getting round this mental block by inviting researchers to start their narrative with the research.

Notwithstanding this expressed preference for a chronological approach to the ordering of the template, there was also feedback to say that academics quite reasonably have far more to say about their research than they do about the subsequent effects of that research in wider social or economic contexts. These Pilot Institutions suggested that preparing a rounded or balanced impact statement required much greater attention to be paid to developing the impact story, and that starting a case study with the impact was important in a symbolic sense; it helped to keep authors focused on the most important and challenging aspect of their story.

Several respondents noted that their units had taken advantage of the invitation to prepare a small number of supplementary cases, and that this was thought to be a good / safe means by which to test the boundaries of the HEFCE Guidance, as the work would be judged by the respective Assessment Panel, but would not form part of the institutional assessment and sub-profiles.

The rules governing the numbers of case studies or lengths of the impact statement were considered to be reasonable by the very great majority of Pilot Institutions. However, several Pilot Institutions commented on the challenge this posed for smaller research groups.

Several Pilot Institutions went as far as to suggest the formula might be better if it were non-linear, such that small groups would not be limited to a single impact case study while large departments would not deluge their respective Assessment Panels with many tens of case studies. One pilot suggested a tapered approach might be more appropriate, so for example, research groups with 20 or fewer staff might use a ratio of 1:5 (but with a minimum of two cases), those in the middle ground, 21-100, would apply that ratio for numbers up to 20 and 1:10 for subsequent staff, and the largest departments, with say 101+ would apply a lower ratio again, say 1:20, for all staff above the threshold.

In our worked example, a tapered approach requires a similar total number of case studies, so it is assessment neutral, but with proportionately more work / cost to the smaller and medium sized units[2]:

- 15 FTEs = three case studies as compared with two presently
- 40 FTEs = six case studies as compared with four presently (4 + 2)
- 200 FTEs = 17 case studies as compared with 20 presently (4 + 8 + 5)

[2] In addition to lightening the workload of the larger units, a tiered formula might also give some small 'quality' advantage to larger departments, because they would have proportionately more researchers and research from which to select their very best case studies. That is, if one assumes there is some structural relationship between the size of a research population and the proportion of 'exceptional' impacts within the distribution of all impacts.

3.4 Assessment criteria

Several respondents expressed uncertainty over the precise meaning of the two principal impact assessment criteria, reach and significance, which were defined in the guidelines thus:

- Reach: how widely felt it was; and
- Significance: how much difference it made to the beneficiaries.

The definitions used in the grading system might very well be a source of confusion with the differing degrees of 'significance' described using phrases more often found in innovation literature: incremental, highly-innovative or ground-breaking impacts. This unit of assessment might be seen as meaningless in many arenas, whether that be changing values and culture or some other social phenomenon, like equity or trust. Similarly, choosing the term 'situations' as the unit of measurement for the 'reach' component of the composite impact metric might have seemed rather alien to many, although one can sympathise with the author's difficulty in finding an agreeable collective noun for all types of benefits and beneficiaries.

The confusion might also relate to the implicit requirement, taken from the use of the word *extent*, to *quantify* each impact where possible, both in terms of the numbers benefiting and the quantum of benefit, where respondents are once again telling HEFCE that they believe these quantities to be unknowable in a majority of cases.

No one commented directly on the five-point scale that will be used to grade each impact case study, although on a related point a minority did express frustration at the decision made during the course of the pilot to provide feedback in the form of an institutional profile for each unit submitted, and without disclosing the individual assessments. There was concern that breaking the link between individual grades and case studies might reduce insight and learning opportunities.

3.5 Threshold criteria

3.5.1 Timing

A minority of respondents expressed concern over the time period in which eligible research was to have been carried out, with the requirement that at least one research output, for example a refereed journal article, having been published in the period since 1993.

Several respondents considered this cut-off point to be too recent, and went on to state that at least a proportion of their unit's most notable, evident impacts having been critically dependent upon research breakthroughs made several decades earlier. Those institutions submitting units with a preponderance of fundamental researchers were most exercised by this design parameter.

A sub-set argued that there was no obvious need for a specified timeframe for the underpinning research, since it is the timing of the impact that is important in this part of the assessment and which should be coterminous with the overall assessment period. In practice, several Pilot Institutions mentioned that they had been able to include impacts linked to research that was being undertaken in the mid-1980s, more than 25 years earlier.

In a minority of cases, however, respondents expressed concern about the equitable treatment of newer research groups as well as those researchers working in emerging fields, which, almost by definition, would have a shorter history and less extensive impact repertoire in absolute terms. For this group, the research window was felt to be perhaps a little too wide.

On balance, it seems worthwhile maintaining a time window for underpinning research as that has an important symbolism as regards gestation periods and, indeed, the volume of research effort that might be necessary in order for specific impacts to reveal themselves. The 15-year term is a reasonable compromise, given the state of the

art as regards institutional records, and extending this to 20 or 30 years might unreasonably favour the larger and long-established research groups.

3.5.2 Institutional contribution

Paragraph 28 caused widespread confusion, in part, it seems, because it cites a list of types of institutional contribution that amount to exclusions, rather than adopting the positivistic approach used elsewhere in the guidance document. Perhaps more significant is the apparent contradiction between several of these points and other clauses elsewhere in the guidelines (e.g. membership of expert advisory groups or high-level committees, the impact from which can be included where the advice and outcomes were dependent to a significant degree upon the research results of the individual in question).

There were questions as regards the meaning of the two types of research-based activity, which is to say (i) research proper and (ii) advice to third parties that is based to a substantial degree on that research. There was also confusion over the extent to which it was a requirement – a threshold criterion – for an eligible impact to be linked to research carried out within the institution *and* to subsequent efforts by the institution to apply or otherwise exploit the outputs of that research.

3.5.3 Quality of underpinning research

Several respondents expressed concern over the definition of research quality and more particularly the specification of a practicable means by which to determine consistently and unambiguously whether underpinning research was likely to be judged to be of sufficient quality (by an Assessment Panel). In one case, a respondent applauded HEFCE's advice on this issue as being agreeably simple, which was to check that the research in question was associated with a research-council grant or had led to peer-reviewed outputs.

3.5.4 Evidence

Pilot Institutions struggled with the meaning of several other key terms, in particular what was meant by evidence. Several Pilot Institutions suggested they had taken a narrow view, evidence in a semi-legal sense, which is to say, material that would stand as proof of work done (a bibliographic reference) or an impact realised (a named referee). Others took a more inclusive view, which is to say any *thing* – object, fact, testimonial, anecdote, etc – that might prove helpful to a reader in forming a judgement about the nature of the given impact and its dependence upon research carried out at the HEI in question. There was also a question posed about the weight of evidence: how much evidence is likely to be considered enough? This was a particular concern in those areas where evidence – relevant, objective facts – of impact was much harder to obtain or even conceive, like cultural impact. Where Pilot Institutions provided substantive written feedback or oral evidence on a specific point, those points have been paraphrased and presented as anonymous quotations, indented and italicised as follows.

> *The meaning of reach and significance was problematic, and just how much information / evidence was required (in the impact statement and case studies).*
>
> *Impact can be realised through the cessation of some practice or other, where research has shown a policy or process to be problematic or otherwise undesirable, so how does one go about counting the benefits of not doing something?*

Several Pilot Institutions asked how far it might be reasonable to expect their research groups to go in order to obtain evidence of important yet intangible impacts, which links with a related concern about the meaning of 'Interim Outcomes' and 'Eventual Impacts.'

In its opening pages, the HEFCE Guidelines include a critical section discussing the definition of socio-economic impact where an Interim Outcome or impact is described in somewhat abstract terms as some *thing* that has happened as a result in part at least

of underpinning research carried out within the HEI, and is likely to be an important milestone, possibly one of several, that one might find along the journey from research to the Eventual Impact. The introduction of Interim Outcomes is rather important, as it is an open acknowledgement that many of the most dramatic science-enabled impacts might be realised very many decades after the original research breakthrough was made and, as such, the time-lags are beyond the scope of an institutional assessment system such as REF. These interim or intermediate outcomes are equally important for a majority of research where the ultimate impacts on our well-being, health, environment or economy are going to be terribly diffuse and essentially impossible to distinguish from other areas of endeavour or connect back to specific programmes of work. The section takes the notion of Pathways to Impact, echoing the work of the research councils, wherein good research, whether fundamental or applied, can be expected to influence, in some degree at least, the knowledge, assumptions and actions of others, with that effect assumed to radiate outwards from one's immediate partners and peers to the wider academic community, while also helping shape understanding and values in the wider community, directly through education and outreach and indirectly through knowledge spillovers. The Pathway to Impact model assumes that in at least a proportion of cases, publicly funded academic research will lead to formal outputs – a paper, a book, an exhibition, or whatever – that will be consumed by non-academics, whereby the resulting insight might lead one or more of those parties to change their views or behaviour, an (interim) outcome of the research where, in the fullness of time, that new understanding might provide the platform for the development and implementation of a new service, therapy, piece of legislation and so on, that proves to be beneficial to some segment of the population, several points removed from the original research, an impact.

3.6 Subject-specific issues

When discussing the appropriateness of the HEFCE Guidance, the principal, subject-specific issues related to the:

- Absence of case examples or terminology that would be recognised by different disciplines
- Inclusion of a sample list of impact indicators that was somewhat confusing as regards its purpose and off-putting due to the incomplete and stylised menu

On the first point, several Pilot Institutions suggested that future guidelines might prove to be more effective at communicating their key messages if the explanation of the generic process and rules was to be interleaved with at least some subject-specific material.

On the second point, the Guidance included an annex (Annex G) that set out a *draft 'common menu' of impact indicators*, a provisional list that included a typology of nine broad classes of socio-economic impact, which ranged from delivering highly skilled people to more innovative businesses to improved public services and improved health outcomes. Each class of impact was linked with a set of performance dimensions, or indicators, including for example, the following:

- Staff movement between academia and industry
- Increased turnover / reduced costs for particular businesses / industry
- Influence on public policy debate
- Improved educational attainment among disadvantaged groups
- Enriched appreciation of heritage or culture

The most heartfelt challenge related to a perceived bias within the guidelines more generally in favour of the applied sciences and technological innovation leading to economic impacts. This anxiety seems to be a long way from reality, when one looks more closely at the menu of indicators. However, the ninth *type* of socio-economic

impact – 'other quality of life benefits' – was criticised by Pilot Institutions submitting English returns for the absence of specific indicators, which was seen as symptomatic of an intrinsic bias as well as being unhelpful. It left this particular group with a substantial additional burden as they attempted to work out what kinds of performance dimensions might be deemed appropriate by HEFCE and the English Assessment Panel.

> *The list was too limited in extent. It was biased towards economic impacts linked to the applied sciences by way of technological breakthroughs and associated innovations. Other forms of impact more prevalent within the social sciences or humanities were not addressed to the same extent or with the same confidence. Several Pilot Institutions stated that they felt as though they had been left to fend for themselves on this specific matter.*

The most widespread concern was uncertainty around the status of the 'common menu' of indicators: what was its purpose? Was it primarily intended to encourage people to quantify outcomes to a greater extent than they might otherwise have done, providing a list of suggestions just to get respondents thinking along the right lines? Or were Pilot Institutions required to use at least some of these metrics? Was it version 1.0 of a taxonomy of socio-economic impacts with associated standard / common indicators, which all HEIs will be expected to monitor and report against in the fullness of time?

> *The purpose of the list was unclear, and in particular we had difficulty understanding whether HEFCE and the Assessment Panels would be expecting all case studies to use indicators from this 'common menu'.*

On this last point, several respondents liked the idea of a common menu of standard indicators broad enough to cover all subjects, but it is not clear how feasible such a comprehensive system would be. There might be risks too. A practicable metrics system will tend to be kept simple, such as having ten key performance indicators rather than 100. However, ten KPIs would be unduly narrow and prescriptive given the range of types of impact of interest and the spectrum of research activity and subjects that are covered.

None of the respondents suggested different or better metrics, although these might be evident in their individual submissions.

3.7 Future guidance

The guidance proved to be a bit too generic, and coordinators ran into difficulties when trying to explain the notion of impacts to departments and often had to do substantial preparatory work themselves in order to exemplify it.

A significant minority suggested that the inclusion of worked examples would be a great help, and recommended that HEFCE make full use of the case material. In a small number of cases, respondents went as far as to suggest that future guidance should be developed so that the generic material is fully supported by good, relevant case material and impact indicators appropriate to the specific discipline.

3.8 Conclusions and recommendations

The HEFCE Guidance was well received by institutions, in the main, and proved reasonably helpful to Pilot Institutions in briefing staff and preparing submissions. On balance, the written material and advice was commendable given the newness of the exercise and the methodological challenge of drawing out connections between research and socio-economic impacts.

While the HEFCE Guidance proved helpful overall, most Pilot Institutions expressed some minor concerns. Most commonly, people criticised the four headings within the impact statement, which were confusing and seemingly repetitious. Several key concepts were somewhat ambiguous, interim impact as compared with final impact, for example, or 'reach' and 'significance'. Lastly, most Pilot Institutions that made an English submission noted that the 'common menu' of indicators in the appendices to the Guidance was rather light on measures relevant to the social sciences and the humanities.

Overall, people found the guidelines a little too light on detail and that, for the REF proper, HEFCE should be seeking to expand instructions to minimise ambiguity and bring this material to life with worked examples. The feedback points to several areas where further work by HEFCE might be beneficial to any future guidelines:

- Adding a glossary of terms, bringing together all definitional material in a single convenient reference. The key terms for Interim and Eventual Impacts might benefit from a fuller treatment, with reference back to explanatory models and terminology that has good provenance in the wider, global research evaluation literature;
- Redrafting the text in several of the key sections to improve legibility and minimise the risk of misinterpretation and any resultant inefficiencies or quality issues (e.g. the drafting of the impact statement template headings and the addition of several subsidiary prompts);
- Using the pilot impact statements and case studies to develop illustrative material for use throughout the guidelines and in any supporting presentations or communication activities. This should encompass examples of research-impact pathways as well as impact typologies and impact cases;
- Testing the idea of a banded system for determining the number of impact case studies that must be prepared by a given unit, which might comprise a series of perhaps three ratios (number of case studies to researchers), in place of the current 1:10 linear model, where smaller units submit proportionately more cases while larger units submit proportionately fewer; and
- Using the Assessment Panel's feedback and pilot submissions to extend the draft common menu of indicators to encompass fully each of the agreed broad classes of impact (cultural, economic, environmental, political, social, etc), ensuring the indicators address both interim and ultimate impacts. The ambition is to arrive at a generic list that will include impact domains and indicators of relevance to any unit of assessment.

4. Preparation of submissions

4.1 Introduction

The HEFCE Guidance focused on explaining the scope of the data request and the specific information requirements. However, HEFCE chose not to prescribe the particular process by which submissions were prepared.

The guidelines invited Pilot Institutions to submit a range of case studies, from across the breadth of the unit's research activity and for types of impact that were more or less complex to demonstrate (e.g. inter-institutional collaboration). In addition, the Pilot Institutions were invited to submit a small number of 'supplementary' cases that would be evaluated by the Assessment Panel, with feedback provided, but not judged as part of the submission. These optional cases were expected to focus on types of impact that institutions thought important, but judged to be at risk of being excluded either due to ambiguous or overly-restrictive guidelines and threshold criteria.

In addition to the implicit message within the written instructions, HEFCE's briefings encouraged Pilot Institutions to trial different approaches and, in that sense, the pilot set out to explore and test both the types of research impact and the process of assembling the narrative and evidence. The commitment to the former waned somewhat following confirmation of a decision to publish the resultant profiles: at this point, reputational concerns caused people to take a more conservative approach. In both cases, the insight from any experimentation is contingent upon the Assessment Panels' feedback.

In light of this rather exploratory brief, all Pilot Institutions were asked to describe the approach they had taken in preparing their submission on the assumption that a pattern of good practice would emerge along with some critical success factors.

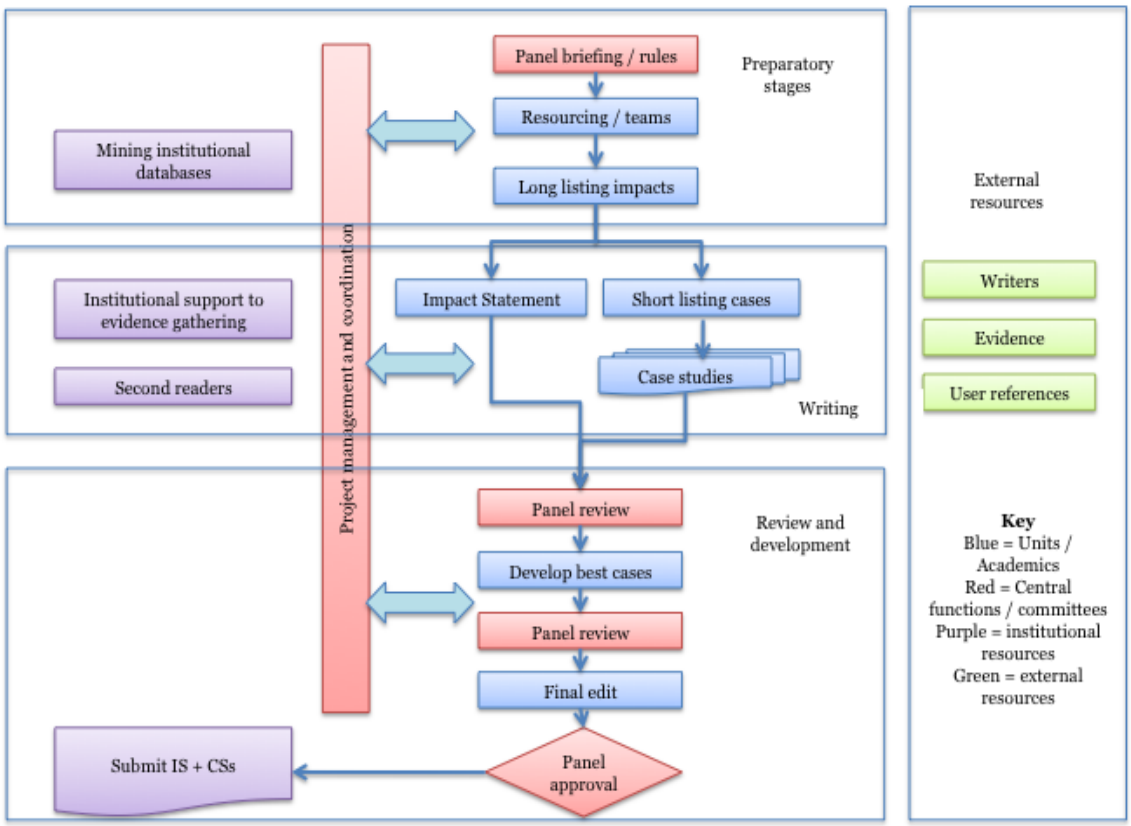
4.2 Submission process

The majority of Pilot Institutions followed a broadly similar approach, in terms of the people involved in preparing the submissions and the process that was followed. In the main, each Pilot Institution implemented the following six functions:

- Project management of the exercise overall
- Researching and writing the impact statement
- Researching and writing individual case studies
- Central (institutional) support with development of drafts and acquisition of evidence
- Obtaining permission for inclusion of external references and other third-party evidence
- Critical review of the drafts and institutional approval

Figure 2 presents a schematic of what appears to have been a typical process, with the individual text boxes naming a process or resource and using a colour code to show which group of actors are taking the lead in that process.

Figure 2 A typical approach to the preparation of an impact submission?



4.2.1 Project management

The exercise was nearly always project managed by a senior administrator from a central service function, typically the research or planning office. The senior administrator takes day-to-day responsibility for the progression of the submission, helping to brief senior officers and faculty on the specific requirements, providing ad hoc advice to the researchers taking the lead in writing content and providing the liaison with HEFCE.

The project manager would typically report to his or her line manager, often the Pro-Vice Chancellor Research, and an oversight committee. He or she had the support of one or more other members of the research administration team, typically with a single, named individual being the first point of contact for each unit of assessment. The central teams also worked with administrative and information services people in the faculties. The size of the central team also reflects the extent to which the institution has established knowledge management systems that include impact-related data.

4.2.2 Oversight of the institutional submission

In almost every case, the exercise was overseen by a high-level committee, which brought together the senior faculty with university directors from specialist functions like enterprise or communications and was typically chaired by the PVC Research. In a proportion of cases, these committees were constituted as formal working groups of the university research committee. The remit varied in extent across the Pilot Institutions. However, three functions appear to be common: engagement of senior faculty; setting of institutional tactics for gathering and packaging evidence; and judge

and jury on content. One might imagine these functions having to be split in the REF proper, for reasons of efficiency, with a single university committee providing direction on strategy and tactics, and a series of smaller working groups carrying out the 'peer review'. One might also expect to see these working groups evolve to include a proportion of external experts, research-aware specialists from user communities, whether that is a technologist from industry, government scientists or policy researchers from the voluntary sector. This combination of knowledge producers and knowledge users on review panels was evident in a small minority of cases, even within the pilot.

4.2.3 Preparation of the submissions by the UOA

The units typically appointed a small group of senior academics to drive the development of their submission, usually with a departmental research director as coordinator, to work across the different schools and research groups and to identify research impacts and coordinate their elaboration in an impact statement and a series of impact case studies.

The two principal dimensions where we saw important differences in approach were in respect to the balance of effort between, on the one hand, central services and the faculty (degree of devolution) and, on the other hand, between the academic leadership and the wider faculty within the units (top down versus bottom up).

> *The submission was coordinated through the central research support office, but the majority of the work was done by the academic staff that carried out the research and brought about the impact. The selection and development of the case studies was led by the research pro-deans for the faculties in which the participating schools were based (coincidentally the pro-deans were also members of the respective schools). They worked closely with the heads of school, research directors and research group leaders. A group of senior staff (primarily research pro-deans), chaired by the pro-vice chancellor for research reviewed the case studies and impact statements and provided feedback to the authors – both on the content of the case studies and on which case studies should be submitted and in effect approved the submission. It quickly became clear that the identification, description and evidencing of impact relied heavily on the knowledge description of the impact and research needed significant academic input.*

> *Academic research co-ordinators of the two departments met with their heads of department and identified a range of possible case studies, soon focusing on six. This team of four then asked for details from the principal investigators in the areas of the studies. They sometimes got short responses that did not give enough detail to write up a case, or long responses that were not necessarily focused on impact as defined by HEFCE. These cases were extensively rewritten by the research coordinators, and then re-written again by the HoDs to give the submission more overall coherence.*

There was a top-down and bottom-up split, one-third to two-thirds, with the choice of approach sometimes tactical or experimental and at other times simply reflecting the size and complexity of the unit in question and the prevailing culture.

> *A different strategic approach was taken with each of the two subject areas: Earth Systems and Environmental Sciences (ESES), and Clinical Medicine (CM). For ESES: A bottom-up approach was followed, with a sharp focus on a single UoA: a single range of impact scenarios were tested and we were less concerned with impact scores. Academics led the process through a steering group and an academic coordinator. Researchers were heavily involved to capture their knowledge of the impact of their research (i.e. academic ownership of process with central R&E office providing mainly supporting data and to review impact statement and research context document). No staff outside the university were used to prepare the submission. For CM: We used a top-down approach, looking at best practice for a wide range of UoAs (19). This time we were concerned with threshold impact scores and developing a worked example to apply across the university built into the university's annual research output monitoring programme. Therefore there was significant central control of the process and use of external contact with third parties to corroborate claims. We were also more selective about choice of joint submission to ensure HEI lead is evident in case.*

> *The Faculty of Medicine took a top-down approach to collating the submission for the pilot. The Faculty of Medicine's Strategic Research Committee identified and prioritised approximately 45 realised impacts (for 40 case studies). The academics involved, with significant support from the faculty's strategic research manager, then compiled the case studies. Further editing and identification of indicators of impact was carried out by the strategic research manager. The Planning Division and an editor employed by the faculty*

for this purpose also provided significant input. The impact statement and research context were written by the faculty's strategic research manager.

The Department of Physics took a bottom-up approach to the pilot. Heads of Group within the Department of Physics put forward 38 examples of realised impacts (for 13 case studies). The Head (and associate Head) of Department, with support from the Planning Division, prioritised the impacts and provided feedback to the academics regarding drafting. The Planning Division and Faculty of Natural Sciences' strategic research manager provided significant input to editing the case studies. The impact statement and research context were written by the head of department.

When the process was top-down, the senior faculty, or even a single senior academic, compiled a list of impacts from their personal knowledge and archives and prioritised the best cases quickly, before going on to draft cases either on their own or, more often, in collaboration with the principal investigator. Even with this top-down approach, the individual researchers ultimately played a central role in the preparation of the individual case studies.

With the bottom-up approach, the UOA project team would typically brief all members of faculty and invite suggestions for potential case studies from all. In several Pilot Institutions, the units went as far as to run surveys to identify candidate case studies. It seems likely that this comprehensive and systematic consultation of all faculties would be preferred to a more top-down approach in the REF proper, albeit informed by data generated from ever-improving corporate repositories.

Typically, it was the principal investigators who went on to write the case studies, with the support of faculty administrators or librarians, usually with the central services team providing an early challenge function, and feedback on the unfolding case study, before the more formal review by the impact assessment committee. In a small minority of cases, members of the central services team interviewed academics and prepared the first drafts of case studies in order to minimise the burden on the faculty, with researchers being invited to correct any factual inaccuracies or otherwise develop the case through further elaboration or new additional evidence. This ghost-writing technique was also implemented to ensure a sharper focus on research outcomes and a more consistent treatment of individual cases and presentation of supporting evidence. The approach does create a bottleneck and resultant time pressures and, in one instance, the centrally managed approach was used to develop a series of exemplary, model case studies before the writing of the remainder of the case studies was devolved fully to the researchers.

The academic contributors were the main source of substantive content in the very great majority of cases, with central support staff providing a challenge function – checking for compliance and legibility – and ultimately carrying out some limited sub-editing and final proofing. The central support team tended to play a more active role in drafting the contextual overview and impact statement.

4.2.4 Institutional support

Setting aside the very important contributions of the individual administrators within the central services departments, most Pilot Institutions looked to support the preparation of their submission through mining archives of press releases, compiling statistics on knowledge transfer and public engagement, running searches on institutional repositories and searching more widely for third-party data, whether those were published statistics or independent evaluations.

It is not clear from the feedback just how much these institutional resources were used, or how critical they were. However, a majority of respondents did go as far as to say that they will be developing their information systems and reporting procedures going forward in order for these corporate repositories to play a fuller role in the future.

4.2.5 External support

In around half of all submitted units, research users were contacted as part of the process of preparing the case studies. In a small minority of cases, research users were invited to comment on both the impact statement and the case studies.

> *An attempt was made to obtain feedback on the Social Work and Social Policy draft submission from a 'user'. However, given the short timescale for the pilot exercise this proved very difficult, especially as the 'user' was not familiar with the proposals for the REF / guidelines for assessment of 'impact'.*

> *Where possible we tried to engage research users for testimonials and evidence. Overall this proved to be relatively easy. However, there were issues around confidentiality and commercial sensitivities. It was also challenging to ensure that our research users were able to respond within the tight time constraints available.*

> *We used an external panel of readers from industry, the NHS and academia to review some of our Clinical Medicine case studies. We had a small group of four reviewers and they provided extremely helpful feedback on the case studies.*

Feedback from those Pilot Institutions reporting little or no contact with users suggests that this was often the result of practical difficulties, where initial enquiries were made, but to no avail, and other priorities simply crowded out further efforts to track down external people.

In a good proportion of the rest, Pilot Institutions took a tactical decision not to make contact with research users, electing to rely upon existing institutional knowledge, at least for the pilot. This was partly about the pressure of time, with Pilot Institutions choosing to avoid the cost and complexity of an external consultation when faced the need to prepare good submissions in a four-month period (November to February). Additionally, it appears to be a function of the size of the submitting units, where the smaller (<30 FTEs) and more homogenous units (single departments) were more confident in their ability to come forward with the required number of very good case studies. In a very small minority of instances, the Pilot Institutions had established knowledge management systems that were good enough to inform the compilation of long lists of candidate impacts and to provide a good start for impact stories and quantification.

Where Pilot Institutions did make contact with research users, a majority stated that it was a rather time-consuming and uncertain process, and when scanning all of the returns, people encountered several recurrent challenges:

- Individual research users had moved on in some instances, and this tended to weaken the organisational memory of the links between a specific piece of research and some wider improvement
- Where research users were still in post, people were content to talk about the importance of the research in general terms. However, it was difficult to quantify the nature or extent of the resultant changes
- Where relevant data did exist there were often commercial confidentiality or other data protection issues, which hindered access

Just three Pilot Institutions commissioned external people to help prepare parts of their submission, albeit in two cases this amounted to contracting in former members of faculty for additional capacity to help with writing and editing. These individuals were familiar with past work and were well known to the current staff. In only one instance did a pilot commission a consultant to more formally research the impacts of one of its case studies, and while this was considered to have been a valuable exercise, it was concluded that the cost of such investigations (a loss-leader for the consultant on this occasion) would be prohibitive for deployment on a more general basis, within the REF proper.

> *It was decided in discussions with the Physics Department to focus attention on making at least one of the cases a substantive, comprehensive and fully evidenced study with an historical perspective and a global analysis of influence (the Organic Electronics case, which highlights one of the most important areas of research to come out of the Physics*

Department). In addition to a committed effort internally, the coordinator also commissioned (c. £5K) a consultant to research global influence through, amongst other things, citation analyses of the central patents. Subsequent to the REF exercise, this analysis has proved very useful to the department in non-REF activities and the exercise can be considered to have been value for money. It is unlikely, however, that we would follow the process as it stands for the actual REF.

Elsewhere, Pilot Institutions found they had been able to make good use of impact case studies prepared by independent analysts and commissioned by the research councils, in particular the AHRC and ERSC, both of which have devoted substantial effort in recent years to developing this particular evaluation methodology.

4.2.6 Panel review

The Pilot Institutions almost always constituted a high-level group of university officers and faculties to critique the draft impact statement and all of the case studies, providing overarching and specific feedback on the development and drafting of the submission.

The majority of respondents stated that the final statements and case studies had required three or four iterations of writing, critique and editing, before the committee was ready to approve the final submission. The number of iterations and re-work should reduce in time as everyone becomes familiar with the notion of research impact and how best to describe it. However, it seems unlikely, for the foreseeable future at least, that anyone will routinise or semi-automate the act of connecting research with specific impacts.

We set up an Impact Task Group chaired by the PVC Research and Enterprise, with representation from academic staff, research support officer, enterprise support, and corporate communications, to oversee the process and decide on the case studies to be submitted.

A Steering Group, led by the pro-vice chancellor for engagement on behalf of the university research committee, met at the start of the project to gain an understanding of the requirements and process. A submissions review group (comprising members of staff from the submitting schools, the university's central public relations team, the research and consultancy division, the planning division and the director of the strategic development directorate) met a fortnight prior to the submission deadline to review the style and content of the submissions. Final approval for each submission was provided by each head of school and the pro-vice chancellor for engagement.

At university level, an REF steering group, comprising the PVC research, senior academics with knowledge of RAE2008 and the Head of Research Support, coordinated the university submission and worked with faculty associate deans for research and individuals at department level. The steering group set the timetable and reviewed draft submissions, providing feedback on the case studies and the impact statements to the departmental REF coordinators. The steering group approved the final version of the submissions.

4.3 Subject-specific issues

There were very few subject-specific issues reported, as regards the submission process.

There were several instances where Pilot Institutions reported different approaches, in the detail at least, having been adopted by their respective units of assessment. However, there is no obvious pattern in the data that would suggest earth scientists might need to follow one convention while physicists must prefer an alternative. Indeed, there are as many variants in approach evident amongst the submissions from units within the same subject area as there are across subjects. The approach adopted appears to be more likely to reflect the size and structural homogeneity of the submitting unit, rather than epistemic differences amongst subjects. Equally, a significant minority of Pilot Institutions stated that they consciously implemented a standard approach for both submissions, and none of these organisations reported this design choice as having been problematic.

Looking across the 29 Pilot Institutions, however, one possible subject-specific pattern did reveal itself: the majority of the 13 English departments adopted a bottom-up

approach to the identification of impacts involving the whole of the unit in the discussion of candidates and looking to the individual PIs to take a lead in writing the case studies. There was a more even split between top-down and bottom-up approaches in each of the other four subject areas.

Two Pilot Institutions reported that their English groups had devoted more effort to conversations with research 'users' in order to elaborate specific impacts where academics in their second submitting unit had not felt the need, within the context of a pilot exercise at least, to systematically contact research users. It was said that the latter groups had an established view of specific research impacts sufficient to the requirements of the exercise.

For the pilot exercise at least, it seems that the approach adopted had little to do with differences in the nature and locus of research impact for a given subject.

4.4 Wider initiatives

In around one half of all Pilot Institutions, participants elected to use their participation in the pilot exercise as the basis for a wider initiative to explore research impact. In several institutions, these shadow exercises were implemented across the university or college, with all units of assessment being fully briefed on the REF pilot and then being required to make a (partial) mock submission, with impact case studies being submitted to the impact Assessment Panel for careful critique and feedback. In one Pilot Institution, the impact case study template has been added to the information requirements of the annual, institutional assessment of departmental research efforts.

> *We used the pilot as a prompt to carry out a much more broad-based assessment of research impact, as part of our more general preparations for getting ready for REF, with every one of our 24 UOAs being required to prepare one impact case study as part of our university-wide annual research assessment. The case studies were submitted to the Impact Review Panel for consideration, and will be given feedback in due course in a similar vein to pilot case studies.*

> *All other departments within the college were asked to produce a selection of impact case studies, based on the same format, criteria and guidance as that used in the impact pilot study.*

> *We performed a mock-REF exercise immediately following the pilot, for which impact statements and case studies have been collected. Lessons learned from participating in the pilot exercise have enabled us to work with members of staff across the university and provide tailored support to ensure the timely preparation of impact case studies.*

> *We used our participation in the pilot to initiate a shadow pilot exercise in a further two UoAs: English and Applied Social Studies. This exercise remains at an early stage and we are currently exploring our understanding of research impact in the context of Arts and Social Sciences research with academic colleagues. The intention is to develop detailed case studies and impact statements following receipt of feedback from the pilot exercise.*

> *The Physics pilot was used as a basis for a faculty-wide impact pilot in the faculty of science and technology*

Elsewhere, several Pilot Institutions used the exercise as an opportunity to test their institutional resources and knowledge management procedures: what existing resources did they have that might help them to more systematically develop preliminary lists of research impacts and inform their appraisal and subsequent development. The general answer appears to have been that there were no good corporate systems, institutional repositories comprising academic data and so press releases / news stories tended to focus on topical issues about future potential.

> *We used the pilot as an opportunity to try to test whether we had any existing corporate sources for impacts in the two UOAs, and we started an analysis of our database of research news stories, intending to review what had appeared on the university website over the last five years. However, the very tight timetable for the pilot curtailed this. Our brief review of the research news stories suggests that these would not provide a systematic data source, but more work could be done to follow this up. Involvement in the pilot did, however, help the impact working group in setting up and running a programme of successful faculty-based impact workshops and presentations, and we will be continuing these in the coming months.*

A significant minority of respondents suggested they had used the pilot as a prompt to begin a wider conversation about research impact across the institution, and specifically what is meant by socio-economic impact and where and how people might do more to facilitate impact in the future.

4.5 Conclusions and recommendations

Each of the 29 Pilot Institutions was able to prepare a compliant submission or submissions, with no expressed concerns about the completeness or quality of the material provided, which suggests that a generic model is an entirely practicable solution for the wider HE community.

While HEFCE chose not to prescribe the process by which submissions should be prepared, a standard approach did emerge quite naturally, with central administrators project managing the exercise, senior academics taking the lead in drafting the statements and case studies and high-level committees peer reviewing material and providing tactical advice.

There were variations in the way things were organised across institutions and units of assessment, with differences evident on two dimensions:

- The division of labour between the central team of (research) administrators and the academics. A senior research administrator and central support team took the lead in preparing the submissions in a significant minority of the Pilot Institutions, preparing preliminary drafts of impact statements and case studies based on close interaction with the units lead academics, interrogation of institutional and departmental archives and interviews with the principal investigators. By contrast, in a majority of cases, the central administrators kept more closely to project management and critiquing draft material prepared by academics
- The extent of involvement of all members of the faculty. In several Pilot Institutions, one or two senior academics volunteered to carry out the great majority of the work involved in preparing the impact statement and case studies, selecting cases and drafting text with only minimal support from the principal investigators responsible for a given impact. This top-down approach appeared to be more common amongst those Pilot Institutions with smaller units. Elsewhere, Pilot Institutions adopted a more bottom-up approach within their departments, sometimes running formal surveys of all academic staff to come up with candidate impacts, and inviting individual PIs to lead in the drafting of specific case studies
- At this interim stage in the pilot exercise, Pilot Institutions were equivocal about the best approach. However, people were settling to the idea that there needed to be a strong and substantial contribution by senior research administrators, and their support staff, to help minimise the burden on key academic staff, and a faculty-wide input to the long-listing and selection of impacts and a more cooperative approach to drafting. A minority sounded a note of caution about the potential temptation to delegate rather too much responsibility for preparing submissions to marketing and communications specialists. Professional writers, in their desire to produce stronger and more coherent stories, might unwittingly overstate the nature and extent of the impacts attributable to the research in question. By contrast, other respondents noted that academics often displayed a natural 'modesty' about the impacts they have generated and that many academic authors might benefit from constructive advice from central communications people. The issue seems to be about balance, with a strong sense that the academics must be the final arbiter

There were few evident strong differences in approach, which were subject-specific. However, the English units tended to adopt a bottom-up approach more often than not, where there was a more even split elsewhere. Moreover, the English units appear to have devoted greater effort to making contact with research users.

Almost half of the Pilot Institutions took the opportunity to run a wider programme of internal events and workshops to help raise awareness about research impact and a significant minority went as far as to run shadow impact assessment processes for other subject areas.

We recommend HEFCE move forward with its current model, which is to say inviting HEIs to make a series of generic submissions for evaluation by a subject-specific Assessment Panel.

5. Challenges overcome

5.1 Gathering evidence

The Pilot Institutions were asked if they had encountered any specific difficulties in preparing the submission. Where they reported difficulties, the institutions were asked to explain the nature of the challenges and how they had overcome them.

A significant majority reported that they had difficulty explaining to their academic teams how to identify what impacts their case studies had, and that their teams had struggled to gather evidence describing specific impacts, in particular objective data to help convey the nature and extent of the benefits realised as a result of their institution's research. There were also observations made that the unit of analysis might be problematic, wherein, in the main, the achievements of a single researcher, or small group, are unlikely to produce benefits at a scale that might reveal itself in national or regional statistics, and that quantification might always require primary research, which seems impractical.

Most commentators suggested there were fewer difficulties in proving their research had played a role in the realisation of a particular impact, although several did suggest that attribution was a major headache in areas where many researchers in different institutions had been working together over many years.

5.2 Explaining impact

A small majority cited the openness and ambiguity of the HEFCE guidelines and, in particular, with respect to the *meaning* of research impact, as having caused difficulties in the time and effort taken to arrive at a point where researchers felt confident they could come forward with relevant suggestions and would know how to develop suitable case studies. Institutions mostly overcame this challenge through persistence in ongoing communication and numerous bilateral discussions around the development of individual case studies.

More purposefully, several institutions elected to develop their own case examples, with the senior research managers working closely with senior academics to compile a selection of exemplary, real-world examples that the individual departments would recognise. In two Pilot Institutions, people reported having developed several impact case studies as reference cases to give PIs the confidence to write. In one other pilot, we heard that some people 'got it' immediately and quickly prepared really excellent case studies, and this provided a great platform for some peer learning.

> *Establishing whether case studies met the criteria for inclusion: This was a challenge in some cases, partly as the criteria themselves were not fully defined and part of the purpose of the pilot was to elucidate them.*
>
> *Interpreting whether a potential case study demonstrated 'actual' impact that could be evidenced or was, in HEFCE's emerging impact language an interim impact or only an activity that might lead to future impact required fairly forensic examination. As a general rule, it was harder to produce robust evidence of how an impact had actually changed people's lives than it was to evidence interim stages.*

A small minority cited the pilot's rather linear and instrumental notion of research impact as having been something of a turnoff to academics, many of which felt the chosen model missed the realities of their particular world. The response here appears

to have been to keep talking, while looking for examples of more instrumental impacts, which were quickly adopted and used to secure wider commitment, and to accept some softening of the HEFCE definitions, to more fully accommodate colleagues' view of research impact. This might yet prove to be problematic if, for example, Pilot Institutions have allowed researchers to focus on near-at-hand activity, such as an appearance on a radio programme or the design of a public exhibition, as while such outreach activities might be prestigious, can they really be said to be a reasonable proxy for a changed outlook or new understanding amongst a specific group of individuals?

A small minority mentioned the challenge of persuading academics to contemplate the ultimate consequences of their work, which some consider to be idle speculation while others somewhat diffidently express unease at the idea of laying claim to social impacts that are almost certain to have arisen as a result of the efforts of many actors rather than one. This is more than modesty and points to a widely held academic view as regards the cumulative and incremental nature of intellectual advances: as Isaac Newton famously wrote, 'If I have seen a little further it is by standing on the shoulders of Giants.' This notion of multiple agents is common in the evaluation literature too, and drives interest in concepts like attribution and, more narrowly, additionality.

A small minority commented on the word limit for the case studies, suggesting that it was perhaps too restrictive given the sometimes rather complex pathways and extensive benefits. It was also thought to favour a particular kind of writing style, a skill that has little to do with the questions at issue. This was offered as a comment, but elsewhere respondents reported having to work hard, and through many editorial interactions and re-writes, to arrive at case studies that were to the point. The role of the central and faculty coordinators and the Assessment Panels was clearly key in helping to develop pithy and relevant case material.

> *Keeping to word counts: Succinctly describing an impact and giving its context in only 500 words for a mixed academic / user Assessment Panel was a challenge requiring specific writing skills. Potential case studies were submitted in outline by the academic originators but nearly all had to be re-written by the UOA coordinators, after review and with advice on style from the project teams.*

5.3 Issues with the continuity of research project teams

A small minority mentioned the challenge of compiling a strong portfolio of case studies due to researchers moving institutions as they progress their careers: even with typically low levels of staff turnover, impact timeframes measured in decades can be problematic. It was suggested that it was harder to secure the support of former colleagues when developing case studies. The practical response to this 'leavers' problem', taken by several Pilot Institutions, was to add an additional criterion, wherein a unit only considered candidate case studies where the PI was still in post.

> *Other than a mixed level of response from in-house researchers to provide outline data on potential case studies, the main challenge was around researchers who have moved on to other universities and who had no incentive to provide necessary information, especially as the nature of the data required was often opaque. We experienced a variety of responses: a feeling of a conflict of loyalty to old and new universities (even though the guidelines are clear on this); a sheer reluctance to find the time to go over old research; a complete ignoring of multiple approaches.*

> *There was a perceived reluctance to use cases based on the research of academics who were at the institution during the period but who have since left. Such cases would clearly be more difficult to write than those where the academic involved is still present, but as impact, unlike outputs, stays with the institution, important cases might be omitted from the REF nationally unless these difficulties are overcome.*

A very small minority stated that the timetable had been rather too short, and that this had required people to be a little more expedient than should be the cases in the REF proper, and having had to omit certain potentially good cases or otherwise truncate the development of case studies.

5.4 Subject-specific challenges

Several subject-specific challenges were reported.

English would appear to confront a particular challenge as regards the assessment of impact, where feedback suggests that most *impacts* are likely to be of a conceptual nature, with intellectual advances contributing in a rather incremental, diffuse and non-specific manner to the evolution of our knowledge, values and institutions.

> *There was a difficulty for English in demonstrating breadth of impact across the unit, as research in the humanities tends to be individual rather than undertaken in groups. This challenge was overcome by taking the department's research themes as the starting point and then considering the cumulative impacts from several individuals' work.*

> *Identifying discrete impacts was problematic for English. The process was to start with the department's main research themes and then think about the impacts that had emerged from research activity. The terminology in the guidance seemed more suited to the sciences than the arts and humanities. Academics tried to engage with the terminology by applying their own definitions of impact: Physics felt the time lag going back to underpinning research was about right, whereas English felt it needed to be longer. There were commercial confidentiality issues for Physics, which did not arise for English.*

Connecting specific pieces of research to specific social or cultural developments was thought to be particularly challenging – too many straight lines – and there was a suggestion that English, possibly the humanities more generally, might have to focus on interim impacts to a very large degree. However, even these more immediate outcomes were thought to be hard to evidence, often involving private individuals and voluntary sector organisations with no inclination or requirement to measure 'improvements'. Even engagement with more mainstream media is said to be rather uncertain, as to whether institutions might reasonably expect access to audience figures and ratings, given the commercial sensitivity of the same. There were concerns expressed about the value of media statistics to researchers: what might one infer about research impact from the exit poll of one lecture given at a science festival?

There was also a suggestion that the tendency of research in the humanities to follow the single scholar approach had brought its own difficulties in that their work can be rather long-term and narrowly based and therefore the ratio of case studies to staff might be insufficient to convey the breadth and weight of an entire UOA's socio-economic contributions. The institutions got round this by making case studies of research areas or themes, involving multiple researchers over time.

Perhaps rather unexpectedly, several respondents stated that their English departments were rather better informed about research impact and ready to showcase their contributions to non-academic communities than were their colleagues in other subject areas.

Physicists had some particular concerns about the challenge of assessing impact in the manner foreseen by HEFCE.

> *1. The difficulty in attributing UoA-specific 'ownership' of impact that arises from the work of a large collaboration, either experimental (in our case, a ten-institute international collaboration working on experiments at CERN) or theoretical (an eight-institute collaboration using High Performance Computing facilities).*

> *2. The blurry distinction between the 'use of' new high performance computers and the 'stimulation of' the development of these machines. At what point does close collaboration with a company such as IBM in the development of teraflop or petaflop computers become 'impact'?*

> *4. Commercial confidentiality became an issue. Even in this pilot exercise, and making use of the 'Confidential' tag on the submissions, we still had to remove some of the strongest material concerning cooperation with industrial companies because of confidentiality agreements. In REF itself, this would be a huge problem. It should be noted that this also contributed to the Institute of Physics giving up on a recent exercise to try to quantify the impact of Physics.*

> *It is impossible to quantify the impact of public understanding of science initiatives such as TV and radio appearances, popular lectures, books, etc at the scale of individuals or groups. Across the community, such work clearly has an important impact, but this cannot be sensibly quantified.*

Several respondents noted that research impact was most readily observed and measured where the distance between the two realms, research and impact, was relatively short and the pathways rather direct: for example, applied scientists working in collaboration with a private business seeking expressly to translate the generic insights into commercial advantage.

The socio-economic impact of more fundamental research might go unnoticed and unremarked most of the time, reflecting the diffuse, cumulative and rather unpredictable nature of intellectual advances. Indeed, one might imagine that the occasion where a fundamental breakthrough produces a relatively immediate and direct social improvement is rare, even anomalous. In practical terms, this means the impact case studies track back to particular types of people and types of work. It might also mean that some of the most consequential socio-economic contributions will be excluded by virtue of their rather indirect link with cutting-edge research.

There was a suggestion that institutions had taken the low hanging fruit, and that in some sense the portfolio of case studies was not a good representation of the breadth of research undertaken. This sentiment appears to be rooted in a general sense that impact is most obvious in some narrow areas, within subjects rather than between them, and largely unknowable, at least in practical terms, in most instances. There were also suggestions made that a proportion of the more significant, evident impacts was linked to people and activities outside the very best academic research, albeit dependent upon that work, and which would be overlooked as a result.

> *There is always going to be an over-reliance on a small number of sub-groups, even individuals, whose research happens to lend itself to non-academic impact. The impact submission is not necessarily representative of the UoA as a whole.*

A significant minority of respondents commented on the degree to which a given unit's impacts might be skewed in favour of certain sub-fields or even individuals.

> *One of the major problems we faced was linking these impacts back to single published research papers of a given 'quality'. What we tended to find was that within each department some of the academics do very 'academic' work and focus mainly on research publications while others work in much more applied areas and carry out a lot of 'public / private engagement' activities. The latter researchers can point to multiple impacts from their work, but struggled to attribute those impacts to 'high quality' published papers – the impacts came mainly through their engagement and interaction with the users / beneficiaries and through the transfer of knowledge between individuals on a human level rather than as a result of the users reading academic research papers. Reflecting on this, HEFCE's view of impacts should be more open / flexible and they should be interested in impacts however they are generated.*
>
> *We have 'impact' people who tend not to focus on major publications (monographs etc) because they are too busy doing research and engaging with users, and on the other side they have 'academics' who put most of their effort into research publications and as a result have little time to engage with (non-academic) users. Individual researchers rarely have time to do both well and so tend to specialise in one or other area, however HEFCE's rules require the areas of (best) impact to be related to the (best) publications, even though this is not how things work in reality. Big impacts don't come through published papers but through other forms of engagement!*
>
> *We were left with a strong sense that fundamental academic research, such as that undertaken in this faculty, is much less likely to produce outcomes of the kind envisaged by HEFCE.*
>
> *... the assessment of non-academic impact is not appropriate across all subjects. Indeed, even within disciplines (for example Physics), areas of applied research lend themselves much more readily towards impact assessment than theoretical research. Hence, impact assessment must take into account the subject platform from which the impact has been achieved.*

5.5 Diversity issues

None of the 29 Pilot Institutions cited any specific difficulties around equality or diversity, and more than two-thirds stated categorically that there no gender or age-related issues had arisen in preparing the submission.

5.6 Conclusions and recommendations

The most widely reported challenge was conveying HEFCE's meaning of research impact to academics and, in particular, that it was much broader than economic impact.

The response in most cases was to develop additional illustrative case material and to grasp opportunities where presented, whether that was enthusiastic volunteers or suggestions of self-evidently good and relevant examples of non-academic impacts.

The biggest challenge was the need to acquire evidence to reveal the reach and significance of a given impact. There were many practical issues that stood in the way, and the response of the majority, for the pilot exercise at least, was to firstly focus on the more obvious cases and, secondly, to use whatever narratives, references and statistics that came to hand readily.

The implication is that the individual case studies are not as good as they could have been, had there been more time available to prepare cases and institutions had the resources and could justify greater investment in researching impacts.

Respondents expect the evidentiary challenge to become less problematic in the medium term, as academics acclimatise to these requirements and begin to log evidence routinely. The suggestion is that the Pilot Institutions will all begin to press researchers to do a little more to log interesting developments along the way and track partners and research users – to give a better starting point for picking out the best and most tractable cases – and having scaled the learning curve one can expect people to be a little more focused on developing the impact story.

However, several contributors argued that while new behaviours and systems will improve the evidence base, and the ease with which it can be assembled, intrinsic difficulties will remain, by virtue of the manifold, diffuse and long-run nature of research impact. In light of these observations, we expect many institutions will still select their portfolio of REF case studies in part based on the availability of evidence.

The subject-specific challenges concerned English most often, where impacts tended to be more conceptual than instrumental, so intrinsically difficult to convey and dimension. The response was twofold, to focus on the small number of *obvious* impacts and otherwise to develop case studies that used narrative rather than statistics and relied on self-evidently noteworthy developments or achievements as a proxy for impact.

One respondent noted the particular challenge faced by physicists, who typically work in quite large teams often on multilateral collaborative projects within international programmes of research: this collective quality makes attribution very difficult indeed.

In terms of recommendations, we suggest HEFCE might consider the merits of:

- Developing a communications pack for HEIs to help REF coordinators move their senior faculty and academics along the learning curve, quickly and confidently. It could comprise a range of materials, and advice on their use, designed to help the senior administrators who will oversee the submission process explain the requirements to university officers, senior faculty and individual PIs. The pack should naturally include the HEFCE Guidelines, however it might also include a selection of case studies and even several PowerPoint presentations, to explain the overall exercise and possibly to focus in on critical elements, from understanding the Assessment Panel's expectations to suggestions for gathering evidence. In the interests of quality assuring and streamlining what will always be a quite complex, multi-layered communication exercise, HEFCE might also wish to give some thought to developing a short course for HEI coordinators, which the Pilot Institutions might come together to deliver to their colleagues around the country
- Working with the Pilot Institutions after the fact to develop some advice on light-touch routines to help researchers gather impact-related material continuously rather than in a one-off rush in the last few months before an impact submission is

due. For example, we were told that several institutions are including an impact heading within their standard, staff appraisal template. This might cause people to take note of wider developments as and when these reveal themselves (systematic opportunism), and automatically provide a faculty-wide, digital record of the more noteworthy incidents, which might even be fed into a regular departmental account of interesting outcomes and impacts. Institutions might help people to get the *impact* habit by developing a simple questionnaire – three questions not ten – that anyone could put to their partners, research users or other stakeholders in the six to 12 months following the conclusion of a grant or centre (e.g. do you think you might do anything differently as a result of what we have found here?). Similarly, collaborative agreements and contracts might look to include a simple clause expressing the ambition for all parties to be prepared to support one another's efforts to keep track of developments made possible by their work together

6. Robust account

We were interested in understanding the extent to which Pilot Institutions considered the submission process to have facilitated the preparation of a good and fair account of the impacts attributable to their unit's research.

6.1 A fair and true account

None of the 29 institutions reported that the process had compromised the quality of their account, which is a strongly positive outcome and suggests the arrangements were appropriate to compiling a fair overview of research impacts and a reasonable portfolio of cases.

Several institutions did remark on the timetable having been shorter than they would have ideally liked, while yet others remarked on the pressure of tightening internal resources. However, this did not lead to any feedback suggesting the arrangements had hamstrung submissions.

The feedback reveals widespread self-limitation around the process, which one might interpret as promising more substantive or otherwise better presentations in the REF proper. The fact that this was a pilot exercise, with final specification unknown, led a significant minority of institutions to conclude they should execute the work as expeditiously as they might. In several institutions, senior officers took the decision to limit the preparatory work in some degree, whether that was a decision to focus on those cases where there was a reasonable amount of intelligence at their disposal already or an instruction to not contact research users except where there was a need to obtain approval for the inclusion of a named individual as a reference. The adoption of a more top-down strategy and the reliance on an individual senior academic was perhaps also a function of this being a pilot exercise: it seems likely that once the specification is fully determined, institutions will prefer a more bottom-up approach and, for all but the smallest units, a more collective endeavour.

The most pointed challenges were directed not to the submission process, but to the more fundamental question of assessment of research impact, at least in the manner foreseen by HEFCE for the pilot. This was in part a concern about the evidence base, and the extent to which it would ever be reasonable or affordable to insist that all academic research maintain an ongoing and comprehensive log of its interactions with non-academic communities and the resulting change in understanding, behaviour or strategies. There were concerns about assessing impact too – a familiar refrain about the rather diffuse and conceptual nature of much impact, and even where there is more direct and observable change, these developments are subject to high degrees of uncertainty as to their location, timing and attribution.

6.2 Further issues

On the subsidiary points, respondents picked out the following issues:

- The use of case studies was well regarded by a great majority of Pilot Institutions, inasmuch as it provides a suitable means by which to develop a narrative and illustrate unfolding research impacts, supported by both qualitative and quantitative evidence
- The use of a case study template makes it easier for people writing case studies to refer to previously prepared examples of *good* case studies, and facilitates more consistent submissions, within and between institutions. In marked contrast, one institution remarked on the openness of the template and supporting guidelines and the challenge posed to the university in seeking to prepare a coherent submission (within the units) and facilitate a consistent assessment by panels

- Around half of all respondents found the structure of the headings in the case study template counter-intuitive, and observed that academics prefer to write in chronological order; research first, impact second
- A significant minority of Pilot Institutions suggested the required number of cases should be non-linear. Smaller units really cannot demonstrate the range of work in two or three cases
- Word length limits of case studies were found to be problematic in several cases, particularly where the case study had focused on the work of an entire group or centre, with manifold impacts. The level of aggregation of reporting – individual, group or centre – had bedevilled several Pilot Institutions, both in terms of determining requirements and procedurally in how best to develop, synthesise and present case studies of larger groups or blocks of work
- There were strong advocates of the impact statement, with several respondents arguing that it had been a necessary complement to the case studies, permitting their units to reveal and explain the full extent (breadth and diversity) of socio-economic outcomes and impacts
- There were many detractors, whose criticisms revolved around the effort required to prepare the statement, its ambiguous and repetitious headings and its uncertain value to the assessment process. In two cases, respondents went on to say that a broad and shallow treatment of impact was too superficial and could not easily be judged by an Assessment Panel
- On additional supporting evidence, several respondents remarked upon the difficulties they had experienced in featuring (disclosing) certain notable impacts due to issues of commercial confidentiality and other sensitivities regarding, for example, animal research

6.3 Conclusions and recommendations

The very great majority of Pilot Institutions were satisfied with the quality of the impact accounts they had prepared, within the framework of the HEFCE process.

The impact statement split opinion, as noted elsewhere. However, done well it ought to provide a good platform for wider involvement of faculty and local insight as well as being a more systematic starting point for selecting case studies. Such a phased, sequential approach will be contingent upon the overall timeframe for the submission. Moreover, the apparently high cost of preparing such an overview might very well diminish with subsequent editions, or updates.

The case study approach by contrast was widely endorsed as being the most appropriate methodology with which to illustrate research impacts and research-impact pathways. The case study format is also forgiving of very different types of evidence.

7. Lessons learned

7.1 Introduction

Having provided feedback on the appropriateness of the pilot process and the robustness of the resulting submissions, the Pilot Institutions were asked what kinds of insight they had taken away from the exercise that would be of benefit to the institution.

7.2 Overall

The first point to make is that every one of the 29 Pilot Institutions reported that they had learned some valuable lessons as a result of participating in the pilot. A small minority qualified their remarks, adding that the most important insights would follow feedback on their submission, and that the codification and wider dissemination of the lessons learned would come later.

The great majority was positive about the experience and collectively cited some half-dozen insights and institutional benefits.

The very great majority reported that their institution had a much better grasp of what HEFCE meant by non-academic research impact and just what was going to be involved in order to detail the nature and extent of that phenomenon.

> *Participating in the pilot has been an educational experience that has helped our institution both to understand the likely approach that HEFCE would adopt to assess impact, and to test how it could be applied. One of our main concerns throughout the consideration of impact in the REF has been in the detail of how it could be implemented. The pilot has highlighted the difficulties and pitfalls that are presented in trying to provide such information. We await the results with interest to see what we can learn from our interpretation of the guidelines, and the approach that HEFCE will adopt.*

> *Having a 'trial run' of the impact assessment process was very helpful: it has identified some early difficulties that we can now work to address and provided valuable lessons learnt for other disciplines. However, although the exercise was useful as a dummy run, it is still a major flaw that detailed feedback from the panels on each case study will not be available to Pilot Institutions. A profile alone will tell us little.*

> *Very useful exercise which has helped us to understand more what will be expected in terms of impact. In fact, the university is now thinking of asking every department to produce one case study.*

> *It did cause departments to take a fresh look at their work, and it shed new light on the breadth of good things that have happened in part as a result of the excellent research being carried out. It has helped people to reflect on their wider contributions and produced useful material and insight that will be used elsewhere. The exercise has also revealed how little we do know about the use that is made of our research and equally that where one has a sense of good things happening, we simply don't record that anywhere. Our evidence base is pretty scant.*

7.3 Awareness of research impact

Most Pilot Institutions reported good progress in respect to alerting their academic colleagues as to the importance of more general reflection and recording of the outcomes and impacts of their research, and a significant minority suggested the exercise had demystified the whole concept and won over at least some sceptics.

> *The pilot has helped to change attitudes in a positive way. We feel we have been able to provide some really good case studies and demonstrate some significant impacts and this has helped those in the departments to feel good about what they do and about the pilot itself. There was some initial wariness but this was soon overcome.*

> *On the whole, our participation in the pilot has engendered a more positive attitude towards the assessment of research impact, along with an increased confidence that appropriate examples of impact exist for the disciplines involved in the pilot.*

There is evidence that, amongst academic participants in the pilot, impact has been helpfully 'demystified' and many of the academics involved feel more comfortable with the concepts. Dissemination via our impact workshops has also been very helpful, but it will be a large task to reach the whole academic community.

We hope that staff across the university, particularly within the submitting units, are now more aware of the research impact assessment agenda and we believe that participating in the pilot has helped to increase engagement with the impact agenda throughout the university. Several schools have held impact workshops and some of these have involved a reflection on the experience and lessons learned from our participation in the impact pilot. It is hoped that the pilot has helped to raise the profile and importance of research impact at a planning level, through committees, and that impact will continue to be embedded into the culture of the university.

Participating in the pilot has not changed attitudes a great deal. There are still negative perceptions that it is inimical to blue skies research and will disadvantage smaller universities.

A significant minority expressed surprise and satisfaction at the catalogue of specific research impacts revealed through the process, and the related, windfall-benefit of being able to showcase those benefits to research funders, partners and other stakeholders.

On a positive note, we were pleasantly surprised at the outcome of the case studies. These clearly provided a much broader appreciation of the impact the university's research has had / is having than previously recognised.

The collection of impact stories is something that the university had been planning to do, and therefore involvement in the pilot exercise was valuable in terms of showing us how to collect the information and to start developing a process for recording and tracking it.

There were plaudits too for the cross-institutional and cross-disciplinary learning derived from the pilot briefing events and offline exchanges between participants.

One of the benefits of the pilot was the networking that developed between the Pilot Institutions, which now have a shared understanding of impact. This included the links with HEFCE, which did conduct the pilot in a truly collaborative way and was prepared to change guidance in light of comments. The institution also has a series of case studies that it can use as exemplars for other departments and for other purposes.

The college has benefited from the exchange of ideas and learning from the experience of other HEIs participating in the impact pilot.

On a rather less positive note, the pilot exercise had confirmed the view of all involved that assessing the impact of research is a difficult and substantial task.

What has become clear is that data collection for the REF Impact element will be a major exercise in its own right, both in educating staff in what is required, and in collecting data. The exercise cannot be simply subsumed under the traditional means of data collection, with which staff have grown familiar. It will probably require specialist impact officers, and force researchers to think in different ways about their research. Both are achievable, but how cost effective is this for something that may well only account for <20% of the overall RAE score?

7.4 Institutional information systems

The great majority of Pilot Institutions have concluded that the inclusion of a research impact element within REF will require additional effort on their part to develop departmental and institutional procedures and information systems, better tailored to recording and archiving impact-related data and material.

A minority of institutions reported having access to some basic record of research impacts. However, the systems were rather patchy and while paper archives might provide good historical coverage, the rather more convenient, searchable electronic systems were much newer. For the majority, however, the principal information system was the amalgam of personal files and recollections of their most long-standing academics.

We have traditionally valued impact because of the areas we research and our research funding sources, but the pilot convinced us of the need to track the impacts of our research and that additional resources and systems are needed.

In anticipation of future REF requirements, we are currently developing our institutional repository to capture the impacts of the university's research. We are devising systems to link those impacts to the original funding and underpinning research outputs.

The feedback was rather inconclusive as regards what those new procedures or information systems might look like. This is partly because the Pilot Institutions do not know yet whether they have submitted a sufficient amount of the right kinds of evidence to make their cases. It also reflects uncertainty as to the most practicable and efficient means by which to capture and organise the heterogeneous data typical of research impacts.

We may develop systems that aid academics in tracking the impact of their research, but will wait until publication of the pilot report before taking this further. We are looking at system development in the wider context of tracking impact, taking into account not only the requirements of statutory reporting to other funders, e.g. the research councils, but also the positive promotion of the impact of our research to our stakeholders.

We were already conscious of impact in our work, but the exercise has put pressure on us to document this and has required us to set up forums with our key stakeholders to ensure an auditable evidence trail is available. There is now an ongoing requirement systematically to collect impact data (qualitative and quantitative), and the list of potential impact indicators is a useful framework to think about impact. However, the university has not yet implemented anything systematically at an institutional level to capture 'impact'. The intention is first to receive feedback / scores on the pilot process and then decide how best to do so.

Have no repository for this kind of material and knowledge management, so we devised an intranet resource for the pilot. We have spoken to our software provider about making an 'impact' module to bolt onto our bibliographic and publications repositories, but it turns out to be not quite as simple as it seems as the current databases tend to be very highly structured and link specific individuals with specific outputs. A lot of the impact material doesn't have that neatness, but nonetheless, we are pretty sure the university will have to do something on the systems front going forward.

We are currently considering how administrative divisions can best support the wider research impact agenda, drawing on lessons learned from the pilot. We are encouraging schools to think about the key indicators that will be most relevant to their discipline and how they can systematically collate and store information to support case studies and / or impact statements.

We have implemented a requirement for all members of staff to prepare an impact statement with a case study if possible, as part of the submission to their annual appraisal. This is expected to help staff better understand what is meant by impact and to cause them to tend to prefer more impactful activities rather than less. It might also help to provide a pool of case material to draw upon.

One of the key issues highlighted in the pilot is the difficulty in assessing impact. There is now more focus on collecting any quantitative data on impact as it occurs rather than trying to do this after a period of several / many years.

Internal systems that have been developed are focussed in an Impact Task Group that will continue to collect and assess case studies from across our institution and will lead into a mock REF exercise.

From this commentary, it seems likely that the first developments will be procedural and that most Pilot Institutions will shortly:

- Ask all departments to make a selection of impact case studies, to expand the total institutional catalogue of reference cases and to establish a cadre of people ready and able to prepare such narratives
- Ask all researchers to think more about what good things happen outside academia as a result of their work, inviting people to record any relevant material or data as and when it arises. In some institutions, the key events and data will be recorded centrally through the annual staff appraisal procedure and possibly reported through the annual departmental report
- Ask all departments and research groups to think about any possible impact indicators of generic value to a discipline or sub-field, preferably based on existing surveys and data to facilitate their use on an ad hoc basis

7.5 Wider benefits

The wider benefits are limited at this stage. However, several Pilot Institutions reported that they had implemented a shadow impact assessment exercise (discussed in more detail above) and that this had helped to generate university-wide discussion and had begun to 'warm-up everyone' to the idea that impact might be an interesting and relevant idea. Equally, several Pilot Institutions mentioned their active use of the case studies produced through the pilot exercise in their general institutional marketing activities as well as any specific, relevant funding opportunities.

7.6 Things we would do differently

The very great majority stated that the institution was broadly content with the approach it had adopted and would not have wanted to do things differently, even with the benefit of hindsight. As with most of the other subjects discussed in this report, a minority qualified their remarks with the caveat that their position might change after they had received the results of the assessment process.

A small minority suggested they would do things a little differently next time round. It seems that resourcing was an issue in several Pilot Institutions, with a sense that they had relied too heavily on a small number of individuals and that a small central team – with a significant amount of support from senior research administrators – would have reduced certain stresses and pinch points amongst the very costly, busy senior academic staff. It might have improved quality overall through a tougher and more negotiated approach to the submission.

One institution said that with hindsight it would have gone for a shorter response, more focused on strong research themes. Minimised the use of technical jargon, explained all acronyms and avoided hyperbole.

7.7 Lessons learned

Every one of the 29 Pilot Institutions reported that they had learned some valuable lessons as a result of participating in the pilot. A small minority qualified their remarks, adding that the most important insights would follow feedback on their submission.

The great majority were positive about the experience and collectively cited some half-dozen insights and institutional benefits, wherein institutions had:

- A much better grasp of what is meant by non-academic research impact, and in particular its broad scope
- A much better sense of what will be required in order to detail the nature and extent of research impacts, and how to avoid at least some of the potential pitfalls in preparing submissions
- A new perspective on the many excellent things that have happened as a result of research undertaken within the university

It seems likely that the first developments will be procedural, and that most Pilot Institutions will:

- Ask all of their departments to develop a selection of impact case studies, to expand the total institutional catalogue of reference cases and to help build institutional capacity
- Ask all researchers to think more about what good things happen outside academia as a result of their work, inviting people to record any relevant material or data as and when it arises

The great majority stated that their institution was content with the approach it had adopted and would not have wanted to do things differently, even with the benefit of

hindsight. This suggests that the process description captured through this feedback exercise can be used by HEFCE as a model process for the community overall.

8. Cost to prepare a submission

8.1 Introduction

We asked all Pilot Institutions to provide an estimate of the internal effort involved in preparing their submissions, along with an estimate of the breakdown of effort between the case studies, impact statements and coordination. We also asked about any external costs incurred in preparing the submissions

8.2 How many staff-days did the process require overall?

Twenty-five of the 29 Pilot Institutions supplied us with data about the number of staff-days their submissions took to prepare.

At the most general level, we can say that it took the 25 institutions 1,414 staff-days in total to prepare their entire submissions, or an average of 57 staff-days each. This aggregate view is of limited value, however, because institutions prepared submissions for either one or two units of assessment and those units differed by an order of magnitude in size (e.g. 20 FTEs versus 200 FTEs). These structural differences are reflected in the size of the range of estimates for total staff-days, with the lowest estimate being only 11 staff-days (covering 27 FTEs and involving four case studies) and the highest being 194 staff-days (covering 519 FTEs and 53 case studies).

The total number of units of assessment covered by the 25 institutions was 43, so we have been able to calculate that the average amount of effort involved *per unit of assessment* was 33 staff-days.

The number of FTEs covered by these submissions was 2,700, so the average amount of effort involved in preparing a submission *per FTE* was 0.5 staff-days.

8.3 What was the breakdown of effort by function?

Sixteen of the Pilot Institutions provided us with a breakdown of how their time was split between preparing case studies and preparing impact statements, and six of these also provided an indication of the effort involved in coordinating the exercise.

8.3.1 Effort involved in preparing case studies

The total effort expended on preparing case studies across the 16 institutions providing data was 689 days, or 43 days per institution. The total number of cases produced was 218, which equates to an average of 3.2 staff-days per case study. The case studies covered 27 UoAs with a total of 2,046 FTEs, so the average cost of preparing case studies can therefore also be expressed as 0.3 staff-days per FTE.

The smallest amount of effort per case study for any given institution was 1.5 staff-days and the largest was 18 staff-days. Overall, and quite reasonably, those preparing submissions involving large numbers of case studies were much more likely to have expended a lower amount of effort per case than those that only produced a small number. This is exemplified by the fact that the average amount of effort per case study for those producing ten or fewer cases was 5.9 days, while the average effort per case for those institutions producing more than ten case studies was just 2.7 days. It is therefore likely that within the context of a full REF exercise, the effort involved in preparing each case would be toward the lower end of the estimates provided here.

8.3.2 Effort involved in preparing impact statements

A total of 16 institutions provided data on the number of days taken to prepare their impact statements. The total effort on impact statements across the 16 institutions providing data was 220 days, or 13.7 days per institution. The total number of impact

statements produced was 27, which equates to an average of 8.1 staff-days per impact statement. The impact statements covered a total of 2,046 FTEs, so the average cost of preparing impact statements can therefore also be expressed as 0.1 staff-days per FTE. The smallest amount of effort per impact statement for any given institution was one staff-day and the largest was 23 staff-days.

8.3.3 Effort involved in coordinating the exercise

Just six institutions provided data on the number of days taken to coordinate their submissions. The total effort on coordination across these institutions was 171 days, or 28 days per institution. The institutions produced nine impact statements and 73 case studies – 82 outputs in total – so we can calculate that there were 2.1 staff-days of coordination effort expended on average per 'output'.

It should be noted that the figures have been significantly influenced by a single institution that (indirectly) reported having spent over 100 staff-days on coordination (of 21 case studies and two impact statements). If this 'outlier' were excluded from the analysis then the average coordination effort per institution falls to just 6.2 days (covering seven impact statements and 52 case studies). The coordination effort for these five institutions therefore averaged just 0.6 staff-days per output, or just less than 0.1 staff-days per FTE covered by the exercise.

8.4 Breakdown of effort by discipline

We also looked at the effort involved by discipline, in order to assess whether the submissions in certain subject areas appear to have been more or less resource intensive than in others. However, only a minority of Pilot Institutions provided a breakdown by discipline and it was concluded that the small number of responses and the intrinsic structural differences prohibited any meaningful analysis of possible differences in effort required, from one discipline to another.

8.5 How was the time spent divided amongst staff?

A minority of Pilot Institutions provided an estimate of the breakdown in costs and the division of labour between support staff and academics, but once again there was a pretty broad spread amongst what was a very small number of returns, limiting what one might infer. One institution reported using 12 days of support staff to 45 days of academic staff. Another institution reported a 50:50 split while another reported an 80:20 split, in favour of academic time. The ratio is a function of institutional decisions on the division of labour between the central services functions and the submitting departments. However, even in the 50:50 scenario, there will be a real cost to the academic team.

8.6 What was the cost of the work?

Ten of the Pilot Institutions provided details about the overall costs of preparing their submissions, expressed in financial terms. The utility of the figures provided is weakened somewhat by the application of different costing methodologies. In around half of the estimates, people used salary and overhead as the basis for the calculation where elsewhere respondents stated they had used Full Economic Costing (principles). The average day-rates then vary from £111.11 a day (based on salary) to £600 a day (based on FEC). With hindsight, FEC would seem to be the most appropriate methodology, at least for the academic contributions, as there is a real opportunity cost with respect to some small, but unavoidable, crowding out of research activity, which is charged and paid on this basis.

8.7 External costs

The great majority of Pilot Institutions made no formal use of external experts or other contractors in preparing their submission, and none of the 29 stated that they would expect to buy in external help, capacity or competence, for the REF proper.

Five institutions reported that they had bought in external help, typically in the guise of former academic colleagues being contracted for a few days to help with drafting the submissions or providing a critical second-reader.

In just one pilot, a decision was taken to commission a domain specialist to carry out primary, market research in order to give due weight to one of the institution's most notable commercial successes. Feedback suggests that the £5,000 cost, which was quite low in comparison with, for example, single, impact case studies commissioned by the research councils, of doing this on a more general basis for the REF was going to be prohibitive.

In another instance, a pilot had considered contracting with an economic consultant to make use of the BIS-approved techniques used to estimate the net economic impact attributable to the infrastructural and knowledge investments of the regional development analyses. The decision was taken not to proceed with this commission partly on financial grounds but also due to uncertainty as to how the Assessment Panel might respond, positively or negatively, to the quantification of impact based on a single or very small number of cases.

For the five Pilot Institutions together, the average sum spent was £2,540, with the lowest spend being £500 and the highest £6,000.

8.8 Which aspects were the most time consuming?

A majority of the institutions reported that the impact statements took up a large proportion of the time spent, compared with the case studies, and many, but not all, consider this poor value in light of the weighting they are likely to be accorded by the Assessment Panel.

Other areas where they cited pressures were:

- Collecting the evidence of the impact, such collecting citations and book reviews, was the burden of the work involved
- Iterating the draft case studies, in order to arrive at a case study that properly and succinctly addressed a research impact as defined by the HEFCE guidelines

A couple of institutions found that they spent a great deal of time trying to harmonise their various case studies and overall university 'look'.

Overall, most of the Pilot Institutions took the view that the workload amounted to a 'reasonable additional burden'. There was a general agreement that the ratio of one case study per ten researchers was about right, with only a few institutions commenting they would prefer a ratio more like 1:15 and several others asking HEFCE to consider a tapered scheme rather than the linear model used in the pilot.

Several of the respondents noted that the assessment procedure was an intensive and time-consuming process, but immediately added that they could not think of a way of reducing this burden. One institution commented that the collecting of the evidence was the most labour-intensive part of the process, but that this was essential to the concept of assessing impacts and was therefore essential.

There was also the advice given that because the assessment was new, the academic community would need time to adjust and possibly need some support or education about the assessment process.

9. Subject-specific issues

9.1 Introduction

HEFCE selected five units of assessment, proposed REF UoAs, in which to pilot its proposals for the assessment of research impact:

- Clinical Medicine (addressed ten of 29 Pilot Institutions)
- Physics (11 HEIs)
- Earth Systems and Environmental Sciences (ten HEIs)
- Social Work and Social Policy and Administration (11 HEIs)
- English Language and Literature (13 HEIs)

The five were chosen to cover the spectrum of research types, from basic to applied research, and disciplines, from physical sciences to humanities.

In 26 of the 29 Pilot Institutions, HEFCE invited submissions from two of the five UoAs in order to reveal differences across subjects and to help test the extent to which one might devise a wholly generic approach to assessing research impact.

There was an expectation that a generic approach might to be tailored by panels to account for disciplinary differences in the nature of research and its wider influence or impact.

9.2 Overall

In order to test the idea of a generic approach, the Pilot Institutions implemented broadly similar approaches across their units of assessment, with some variation in the degree to which the preparation of the submission was managed top-down or bottom-up.

Looking across the feedback from 29 institutions and 55 units, it seems that this generic approach has worked reasonably well in all cases, with no evident difficulties experienced in one unit and not in others.

The feedback does suggest that impacts are more apparent and more easily assessed in the more applied areas of research as compared with more fundamental work. This basic-applied split was reported as a factor shaping their impact submissions by at least one Pilot Institution in each of the five UoAs. However, some UOAs are likely to include a greater proportion of applied work than others: Clinical Medicine and Social Work and Social Policy are the obvious candidates within this group of five. This feedback suggests that some units might find it proportionately easier to develop impact statements with a larger number of bigger and more obvious socio-economic impacts, and that HEFCE and its Assessment Panels might need to carry out some kind of subject-specific calibration of research fecundity.

> *Different subjects are more or less applied; hence, some disciplines (or sub-disciplines) lend themselves more readily to impact assessment. It is important that panels take into account the nature of the discipline when assessing impact that has occurred.*
>
> *…it is evident that impact in Clinical Medicine is not the same as impact in English Language and Literature, and impact in these will not be the same as in, say, Philosophy or Pure Mathematics. Impact is very subject-specific and this needs to be recognised by HEFCE from the outset. The list of possible impact factors therefore needs to be as broad and inclusive as possible, especially as regards the humanities. It will be very important that each REF Assessment Panel sets out its own definition and understanding of impact and how it is to be assessed.*
>
> *At first I was sceptical and hesitant about whether or not impact assessment would work for humanities subjects. Although we are aware of the substantial impact humanities research has, I was somewhat daunted by the methodological problems involved in capturing those impacts. My experience of participating in the pilot has led me to think*

that it is very possible to assess impact in humanities research and that it is a necessary thing for us to do. I am much more sanguine about the process after participating in this pilot exercise.

Impact case studies were skewed towards forms of impact that generate audit-quality data (broadcast radio) rather than other more diffuse forms where data is harder to identify and the narrative connection harder to establish (for example reflective and philosophical work that changes public behaviour).

A majority of the challenges cited by the different units were repeated in one or more of the other UOAs. For example, issues around obtaining evidence to dimension impacts or research users that were reluctant to share data due to commercial sensitivities.

Some features did emerge, although analysis of the impact statements and case studies might provide a better platform from which to identify the particularities of a given discipline and the mixture and balance of different sorts of impacts:

- Clinical Medicine: no obvious, subject-specific issues here, indeed the pilot appears to have worked well for this field, with numerous impacts reported across health, policy and economic realms (based on the criteria and timeframes specified);
- Earth Systems and Environmental Science: substantial proportion of theoretical work, feeding into global discussions on for example climate change and policy related to this;
- English Language and Literature: widespread sense that this field of research was much more likely to be influencing pedagogy or public understanding and much less likely to be deliver the sorts of 'breakthroughs' one might see in the physical sciences providing the basis for new products or services or new enterprises, although examples of both did exist. Changing attitudes and improving understanding were thought to be harder to concretise and assess, and there were many suggestions to the effect that evidence would need to focus much more on interim impacts;
- Physics: concern was expressed here about the very high proportion of theoretical work and the distance in time and space between this kind of research and the sorts of socio-economic impacts referred to in the HEFCE Guidance. There was also a comment about the particularly collaborative nature of much Physics research, with people working in large teams on multinational projects, and just how one might deal reasonably with attribution;
- Social Work and Social Policy: just one recurrent issue here which was the highly political nature of Social Policy and the suggestion that some of the very best research, with real potential to improve things on the ground, is passed over.

9.3 Conclusion and recommendations

Looking across the feedback from 29 institutions and 55 units, it seems that this generic approach has worked reasonably well in all Pilot Institutions, with no evident and insurmountable difficulties experienced in one unit and not in others.

The great majority of the subject-specific challenges revealed through the pilot relate to the nature and location of impact, rather than the organisation or sociology of the research, and the use of subject-specific Assessment Panels should permit those particularities to be made explicit such that judgements and grades can be appropriately normalised. Moreover, many of the concerns expressed in one broad discipline seem likely to hold in many.

The following lists a selection of subject-specific concerns, one for each of the five subjects addressed in the pilot:

- Clinical Medicine and commercial confidentiality
- Earth Systems and Environmental Science and the importance of interim impacts

- English and the importance of individual scholarship conducted over a lifetime
- Social Work and Social Policy and the valuation of the negation of existing policies
- Physics and the attribution challenges posed by international mega-projects

It seems likely that the impact statements and case studies will be dominated by more applied research and there might very well be value in the different Assessment Panels developing a subject-specific view as to the kind of impact mix one might expect and the proportions of very high impact cases.

Further work on explaining, exemplifying and instrumenting interim impacts should help to counter some of this natural bias in favour of the more obvious and immediate impacts of highly applied research. We would anticipate a similar need for a collective response to developing case material and assessment techniques that can credibly deal with the challenge of presenting / attributing contributions to collaborative and multi-disciplinary research.

Equally, some aspects of research impact are much less well developed in terms of their measurement and instrumentation: evaluation methodologies and performance statistics are more readily available in the economic realm and similarly health and educational gains are closely monitored and studied. The AHRC has been making great strides in respect to the humanities; however, the state of the art is still a long way behind economic impact assessment. This suggests that certain disciplines might benefit from relatively greater effort being devoted to the development of exemplary material by HEFCE or the community and to the development and calibration of acceptable performance indicators.

10. Changes in attitude and behaviour

We asked the Pilot Institutions to tell us how they thought behaviour might change if the proposed impact assessment procedure were implemented.

A majority of Pilot Institutions foresee some likelihood that the introduction of an impact assessment procedure will cause institutions to move more toward applied research, with a corresponding move away from 'blue skies', basic research.

The responses sat on a spectrum of views, with a small minority with strongly negative views, marking one side of the distribution, which sees socio-economic impact assessment as a diminution of basic research, and at the other end of the scale a significant minority of institutions that see this development as being strongly positive. As one might expect, the majority sit somewhere in the middle of the spectrum and anticipate a degree of risk that will need to be monitored and managed.

On the positive side, several institutions imagined the increased attention to research impact as being beneficial in the following ways:

- The growing interest in the notion of research impact, evident amongst all funders, is making academics think harder about the difference their work might make to other academics and to wider society, and this is beginning to reveal itself across the research lifecycle, in detailed improvements in the design and execution of that work, from conception to dissemination
- The REF impact assessment might very well mark itself out from the wider interest because it is concerned to come to a view on actual impacts, things that have happened or are happening, and from a broad cross-section of people and subjects. That kind of reflection ought to make people more grounded. One university claimed that because of the trial the academics within the institution were already happier to talk about the impacts, adding that they wanted this cultural shift to go further
- In the opinion of several institutions, the focus on impact might help to redress the imbalance one sees within institutions and funding councils between fundamental research on the one hand and 'less-worthy' applied research on the other. The people and groups that specialise in the rather muddier, multi-disciplinary applied research – including translational research – might see their academic status given a bit of a boost
- It might encourage HEIs to be a bit more enthusiastic about focusing on the 'utilitarian' side of research, using special skills and capacities to do more with local businesses and other stakeholders
- In addition to a shift in the way that academics think about their own work, some respondents expected there to be shift in the way that the academics interacted with others. They expected to see more efforts to communicate their work and to publicise and disseminate their research outside the academic community. Some of them also predicted that it would encourage greater interaction between university research and business
- Critically, several respondents saw impact assessment as a powerful means by which to show the world how many important, fantastic things universities are doing

On the downside, there were fears that an increased fascination with non-academic impact might cause people to tend to favour non-academic work. A significant minority expect some or all of the following to arise in the fullness of time:

- A shift in the balance of the composition of UK research overall, with some expansion in what was labelled as more obvious, applied research addressing short-term goals that one might imagine business or government should address themselves, and away from more fundamental and riskier research. Taken to extremes, there is a small risk at least that the UK will see a very substantial change in focus that might lead to a collapse in our enviable international scientific reputation and, ironically, two or three cycles hence, we could arrive at a situation where research proper has been broken and the 'unarguable' benefits of fundamental research on the economy would be lost
- There was also a fear that that impact assessment might favour sub-disciplines that were more applied. Thus, one institution that had a large department in research into basic Physics was worried that this sub-discipline could get crowded out by more applied varieties of Physics
- Another fear that was prevalent was that because this new criteria favours projects that have evident impacts it could precipitate a risk-averse institutional reaction. This could lead to academics suffering a loss of control over their work, as institutions may feel obliged to encourage a move toward more applied research. It was also felt that this might create an institutional tendency to promote work that is attractive and consequential to particular interest groups, but which is of very little intellectual value to the discipline internationally

A good proportion of the majority stated that they thought behaviour would change very little and that the biggest impact would be an increase in bureaucratic overhead. Most were rather sanguine about this, seeing the procedure as simply another cost to doing research. Two institutions, however, signalled a note of caution, wondering whether such a burden might not fall unduly heavily on a small group of individuals with heavy existing commitments. The more opportunistic saw the Impact Assessment procedure as being a relatively low cost means by which to gather together and maintain substantial amounts of the kind of stories and statistics they are always seeking out, to help give due weight to major proposals or discussions with strategic partners.

There were many observations about the required changes to the administrative structure that they felt were either needed or would occur anyway. Many of the institutions thought it would be a good thing if the universities kept clearer and more comprehensive documentary evidence of impacts and that this would increasingly demonstrate to government (HM Treasury) the good sense of supporting a large and dynamic university research system.

10.1 Conclusions and recommendations

The very great majority of Pilot Institutions expect the introduction of the REF impact assessment component to bring some change in behaviour and, looking across the feedback, one can see anticipated changes at a number of levels: individual researchers, institutions and possibly even disciplines. At this stage, however, very few of the Pilot Institutions expect the behavioural changes to be as profound as those brought about by the Research Assessment Exercise, when it was first introduced.

A significant minority was vocal in its support of the introduction of a research impact dimension to REF, expecting this to focus the minds of all staff on doing good, relevant work, as well as perhaps helping to boost the institutional status of researchers and research groups with a more applied focus.

A significant minority also foresaw this kind of exercise as being a valuable source of strategic intelligence, which can be used in institutional marketing, major proposals and their offer to the world more generally, from overseas students to major partners.

The detractors see a similar change in focus, but view it in a negative light, arguing that such an incentive might change the composition of UK research, in particular bringing about a diminution in the total amount of excellent, fundamental research. If

this were to transpire in any significant degree, the reputation of UK science - all disciplines, in principle – would be harmed and the economic benefits of basic research would fall.

There are several important, inbuilt protections, which suggest the impact will not be so profound. These include: the relatively small, anticipated share of REF-derived income tied to impact ratings; the exclusive focus on impacts with evident links back to high-quality research; and the limited number of impacts required to be reported in depth in comparison with the number of person-years of research effort in scope.

The majority confirm this more neutral reading, and have suggested the proposed new arrangements as simply adding to the cost of doing research, and do not expect to see the Impact Assessment as proposed change behaviour fundamentally.

11. Advice to other universities and colleges

We asked the Pilot Institutions involved to tell us what single piece of advice they would give to other Higher Education Institutions, to prepare for the future REF.

The most frequent piece of advice offered to non-pilot institutions was the need to start preparing as soon as possible and even before HEFCE has issued the final specification. In short, the pilot has revealed, or perhaps confirmed, that research impact is not a matter that the great majority of academics would track systematically, much less feel the need to report on it. Moreover, what knowledge does exist is typically personal, embodied in principal investigators in the main and to a lesser extent senior and long-standing colleagues.

Institutions for the most part have only the most partial grasp of wider impacts, typically arrived at in an opportunistic fashion, whether through press interest in a research story or occasionally the enquiries of other research funders. From an institutional perspective, systematically tracking staff's research impact is a pretty new departure that will require a great deal of discussion and debate.

There were a few areas where early action was seen as being essential. The most frequently cited was the perceived need to start collecting evidence of impacts continuously and contemporaneously, with one institution commenting that even a deleted email can be expensive when documenting impacts that stretch years back. Another institution suggested that it was important to start thinking about how to collect evidence from third parties at an early stage because that can be an unpredictable element in documenting impacts. Several institutions recommended choosing possible case studies early and selecting more than would be required so that the very best ones could be ultimately chosen for submission.

A significant minority also suggested that non-pilot institutions should quickly look to implement some form of internal project to review the situation locally and to begin to explain research impact to academic colleagues and to debate options for moving forward. Part of this conversation will be about the REF impact specification and the scope of the exercise and the likely implications for a given institution. However, the pilot exercise has led people to conclude that there is other work that might be done to begin to research past impacts and document these potential case studies from a cross-section of good examples across the spectrum of impact types and subject areas. This kind of local, real-world material might help people to visualise what is required and to reassure them they are likely to have relevant experiences in their CV right now.

Several institutions suggested non-pilot institutions invite volunteer researchers to work through a few examples to get a 'feel' for what is required and also to put together some 'best practice' examples to show other academic colleagues what is meant by 'impact'. One institution even suggested starting with the more difficult cases, so that they might be used to show how even complex research can be captured within a case-study format and communicated in a user-friendly manner.

Contributors offered a series of other very specific suggestions, including:

- Think about the obvious wider benefits that have derived from research in your specialist area, wherever it was undertaken, and then give some thought to the extent to which you or your group's work has led to similar sorts of outcomes or possibly even contributed to key steps in these discipline-specific impacts
- Another institution's advice was for academics not to limit themselves to recent research, but think in terms of the full range of eligible and underpinning research they have done across their career
- Another institution said 'not to panic, we've been pleasantly surprised by the whole exercise [and the amount of impact we revealed]'

There was a lot of advice about the kind of procedural and information systems one might need to implement in order to make a reasonable job of presenting an institution's contributions, while minimising the cost and disruption for academic staff and the intrusion amongst partners and research users.

There seemed to be a general agreement that dealing efficiently with these new requirements would benefit from some additional structures and capacity, centrally and at the departmental level. Several specific suggestions were put forward:

- It would make sense to add 'impact' to the duties of a senior officer and his or her equivalent administrative director. These new requirements will affect everyone, so some central champion and arbitrage is likely to be necessary. The detail is to be determined of course, and is likely to reflect the existing situation. One Pilot Institution suggested that non-pilot institutions really ought to put *someone* in charge of Impact Assessment, immediately, so there is time to learn and build up the evidence base. By contrast, another institution suggested that a more incremental approach was likely to be best. People should be wary of going overboard with new administrative structures and appointments and instead look at the many opportunities they will have to slightly modify or extend existing arrangements and individual assets
- Several institutions thought that it was a good idea to have the impact evidence stored and available centrally, and that archiving 'systems' needed to be developed to capture and consolidate the material sitting with researchers
- Yet another institution suggested that it was important to have someone within the department who might stand as an impact champion and liaison; someone who understands the kinds of impacts the department is generating. Whether or not this person was brought in to do this job or already had a role within the department, it would be important to have an active, senior academic on the ground who had this kind of 'local' knowledge and who could act as a communication channel between busy researchers and central administrators or external evaluators
- Another institution suggested that the pilot had revealed that the state of the art – in terms of tracking and reporting on wider impacts – varied substantially between institutions and across disciplines, and that peer learning through networking at institutional and disciplinary level would be a good way of sharing best practice ideas and helping the less advanced to develop their own ideas faster about what kinds of impacts they believed they were producing

11.1 Conclusions and recommendations

The most widespread advice to non-pilot institutions is that all should begin their preparations for the REF Impact Assessment immediately, or as soon as is practicably possible.

The pilot exercise confirmed to all 29 institutions that the act of gathering evidence on research impact is a largely new endeavour, with little institutional infrastructure available to support the process and a heavy reliance on the personal knowledge of senior academics.

This embodied quality led people to suggest that HEIs should look to do more to begin to write down what is already known about past impacts and, going forward, to explore cost-effective options to encourage the collection of impact-related material and evidence

It proved to be non-trivial for coordinators to explain HEFCE's view of non-academic impacts to their academic audience, and substantial effort was devoted to general communication as well as to bilateral exchanges around specific drafts of impact statements or impact case studies. This communications exercise will be even more extensive beyond the pilot, as it will need to engage all active researchers covering all other units of assessment.

Equally, the Pilot Institutions believe that a significant proportion of all their academics remain somewhat sceptical about their personal responsibilities for wider impacts and that many will view the REF proposals with a mixture of anxiety and antagonism.

12. A final comment

12.1 A well-run exercise

To conclude the gathering of feedback on the submission process, we invited the Pilot Institutions to offer one final comment on any aspect of the submission, whether they wanted to restate an earlier key point or add something new.

The great majority of the Pilot Institutions took the trouble to comment on the good experience they had of participating in the trial. People commented on the HEFCE's openness in running the exercise and complimented the REF Team on its willingness to debate issues, respond to suggestions and consider challenges of one kind or another. A majority of respondents also praised their fellow Pilot Institutions for their willingness to discuss their own experiences and offer helpful advice to others, whether through the briefing events or offline through bilateral exchanges. People wished to underline the importance of this kind of peer learning support when confronting novel topics as slippery as research impact and its assessment.

Many of the institutions commented on how useful the experience had been in terms of producing a new and rather valuable overview of their contributions to the wider world and some really useful case examples to include in their marketing and communications work. There were many comments like 'pleasantly surprised at how much we have contributed'.

Another institution commented on how they now felt as though impact assessments could work for the humanities and the social sciences. Yet another felt the trial had helped them to understand impacts across their institution by collecting and examining case study evidence in a way they had never done before, adding 'we now have a great source of case studies that showcase our research'.

Several institutions were appreciative of the opportunity to pilot different types of impact case studies and approaches in an experimental environment without having to formally add the outliers into the final profile.

12.2 Concerns about the manner of feeding back results

The final comments were not all positive, of course, and the one recurrent criticism concerned the proposed method of feeding back results. People believed that the aggregated profiles would weaken an institution's ability to learn from the experimental approach. Many of the institutions submitting larger numbers of case studies were disappointed that they would not receive case-specific feedback. One institution feared that the exercise would lose a lot of its value if they could not assess the extent to which individual case studies were appropriate to the criteria and guidelines.

The Pilot Institutions had been expecting a detailed explanation of the grade they received for each of their case studies, addressing the issues of relevance, quality and evidence. It was felt there was much to learn, given the openness of the guidelines and the novelty of the topic and the impact assessment criteria. Moreover, a significant minority had consciously submitted cases on the boundaries of the proposed scheme in order to better gauge eligibility and relevance. There was a particular interest in understanding how different treatments had worked across subject areas, and a concern that the profiles might mask insight.

Others expressed concern regarding the proposed plans to *publish* the profiles rather than selected examples of good cases, or cases from across the spectrum, because the press would 'inevitably' draw up a ranked list or league table of some sort. The reputational impact might be very real amongst other funders and partners. Several institutions observed that this decision to publish all profiles had immediately led to a

change in their tactics, and a much more conservative approach to the pilot, focusing on more obvious or safe cases.

12.3 Conclusions and recommendations

Overall, the Pilot Institutions have come away from the exercise feeling much more comfortable with the notion of research impact, and its relevance to their institution, as well as increasing confidence in their ability to document research impact.

People commented favourably on the HEFCE's openness in running the exercise and complimented the REF Team on its willingness to debate issues.

A majority of respondents also praised their fellow Pilot Institutions for their willingness to discuss their own experiences and offer helpful advice to others.

Many institutions had also come to a new understanding of their many and various specific contributions to the wider world, and were pleasantly surprised at the results.

Appendix A Questions for Pilot HEIs

1. **How helpful was the 'REF Impact Pilot Exercise: Guidance on Submissions' and in what ways should this guidance be developed to facilitate the impact element of the REF proper?**
   - Did the eligibility or threshold criteria make sense?
   - Were the 'rules' appropriate, e.g. on the ratio of case studies to research staff?
   - Were there any aspects where more guidance would have been helpful?

2. **What approach did you take in order to prepare your submissions?**
   - How did you identify examples of impacts and select case studies?
   - Who did what in terms of coordinating, developing and approving the submission?
   - What role did your departments / academics play in the process?
   - How did you assure the quality of your submissions?
   - Did you approach research users? To what end?
   - Did you use any other external people to help prepare the submission?
   - Were there any important differences in arrangements across departments?
   - To what extent did you choose case studies to experiment and test the boundaries of the assessment process?
   - Did you use the pilot as a basis for any wider initiatives to explore research impact?

3. **Did you encounter any specific difficulties in preparing the submission and, if so, what were those challenges and how did you overcome them?**
   - Did you experience any difficulties engaging academics, or helping them to understand what was meant by impact?
   - Did you experience any difficulties acquiring evidence to dimension an impact, or to prove your contribution to it?
   - Did you have sufficient candidates to allow cases to be let go, where difficulties arose?
   - Were there any equality or diversity issues?
   - Were there any differences in experiences across departments?
   - How did you overcome these difficulties?

4. **Do you feel the pilot processes and guidance enabled you to produce a sound account of the impacts attributable to your research?**
   - How appropriate was the case-study format to the elaboration of specific impacts?
   - Did the number of cases allowed cover an appropriate range of impacts for departments?
   - What key additional evidence were you able to provide in the impact statement?

5. **What internal lessons have been learned through participating in the pilot?**
    - Has participating in the pilot changed attitudes towards the assessment of research impact?
    - Have you / will you develop any internal systems in response to the pilot?
    - Have there been wider benefits of participating in the pilot?
    - What if anything would you do differently, with the benefit of hindsight?

6. **How burdensome was the task of preparing a submission?**
    - How many staff-days were required overall, and per case study
    - What was the cost of this work? What was the split between statements and case studies?
    - Did the submission require any other outlay and, if so, for what and at what cost (£Ks)?
    - What aspect of the overall process required the most effort?
    - How might HEFCE seek to ensure the burden is proportional in the REF proper?

7. **What if any significant issues emerged that were subject-specific?**

8. **What, if any, changes in attitudes or behaviour do you foresee as a result of implementing this approach to assessing impact in the REF proper?**

9. **What single piece of advice would you give to other HEIs to help them to prepare for the future REF?**

10. **What if any other observation or comment would you like to offer?**

Technopolis Ltd
3 Pavilion Buildings
Brighton BN1 1EE
UK
T +44 1273 204320
F +44 1273 747299
E info@technopolis-group.com
www.technopolis-group.com