USING BETTER EVIDENCE

A REVIEW OF THE SCIENCE REVIEW PROGRAMME

REPORT AND RECOMMENDATIONS

Peter Cleasby

Quantera

for the Government Office for Science

July 2008

CONTENTS

EXECUTIVE SUMMARY

Politicians frequently say that the business of government is about making choices. In recent years there has been renewed emphasis on the importance of basing those choices on clear and publicly accessible evidence. It is a key function of the civil service to provide Ministers with that evidence.

The use of evidence is not confined to supporting decision-making. It alerts policy-makers, regulators and service providers to possible changes in their landscape, leading them to question - or to confirm – the assumptions underpinning what they do. Evidence often generates the need for more evidence before anything approaching a policy decision is taken.

So having "good" evidence – and knowing what to do with it - is vital. This report examines the contribution of the Government's Science Review programme, which since 2003 has been aiming to improve the quality and use of science in government, at a cost of approximately £750,000 in 2007/08.

The principle of independent scrutiny as a means of raising or maintaining standards is strongly supported, inside and outside of government, and the Science Review programme has scored some worthwhile achievements. Yet it needs to do more to adapt to the rapidly changing business environment within government; and there is ambiguity about the purposes and intended beneficiaries of the reviews. This has led to weakened credibility with Departments, loss of traction in taking forward recommendations and, consequently, underperformance by the programme in delivering on its objectives. The reasons are described in my report.

Changes have been made in the light of experience, but more radical change is required if the programme is to add sufficient value in relation to the challenge it seeks to address. To repay the investment incurred in scrutinising science, there now needs to be a concerted effort to demonstrate both the value and the centrality of doing so. Adoption of the recommendations in this report can make a significant contribution, by:

- Improving the quality and relevance of the scrutiny process, based on new and measurable objectives, shared ownership and shared benefits, and in particular by encouraging a continuing and dynamic rather than episodic model of improvement.
- Re-allocating scrutiny resources so that they are focussed on delivering benefits in shorter time rather than on in-depth investigation when the need for it is not compelling.
- Increasing the alignment with other business drivers and performance assurance processes in government, and progressively moving to cross-cutting approaches and systems.

Much of what is recommended in this report can be applied to the work of the other evidence and analysis professions in government. Joining-up the work of the professions is widely recognised as an important goal, and the reform of the Science Review programme offers opportunities to make progress here.

Change is seldom straightforward. Strong leadership will be essential to ensure the aims of the changes remain as the central focus, with a programme management structure to engage key interests across government and beyond.

CONTEXT

Science in government

1. There is widespread, perhaps universal, support for the principle of external scrutiny of the management and use of science by government in discharging its policy development, regulatory and delivery functions. An absence of purposeful, targeted and well-managed scrutiny significantly increases the risks of making judgements or basing decisions on assumptions which may not be optimal or which may be unsound. The evidence for that proposition is well-known and does not require repetition here.

2. The current Science Review programme run by the Government Office for Science (GO-Science) and its predecessor bodies was launched in 2003 as one such scrutiny mechanism. A "quinquennial" review of the programme was scheduled for late 2008, but has been brought forward in the light of concerns about the programme's operation raised by the Committee of Chief Scientific Advisers. Consequently, I was commissioned in March 2008 to assess the programme's effectiveness and to make recommendations for changes if necessary. My terms of reference, the review's governance, and the approach adopted are set out at Annex F.

3. It is right to record that GO-Science has not been waiting for this review before starting to address concerns about the operation of the Science Review programme. Weaknesses in the review structure had already been identified and modifications have been made. The arrival of a new Government Chief Scientific Adviser in January 2008 has acted as a catalyst for change by bringing a fresh perspective. However the most recent reviews – of the Department of Health and of the Food Standards Agency – are not yet complete, and I cannot fairly draw conclusions about the impact of these changes. The first part of my review therefore concentrates on the five completed and published reviews[1].

Science and other forms of evidence

4. Although much of this report focuses on the Science Reviews themselves, science[2] is only one – though a major - part of the evidence base used to inform the business of government. Those responsible for policy, regulation and delivery need access to all types of relevant evidence, including that managed by other professional disciplines, such as economics or statistics. All Departments have to develop and maintain the ability to ask the right questions and to ensure that the evidence used in helping answer them is quality assured (or, at least, understanding its limitations).

5. This poses challenges not just for the science and engineering profession but for all other disciplines represented in government services, particularly

[1] Department for Culture, Media and Sport (DCMS), Department for Environment, Food and Rural Affairs (Defra), Health and Safety Executive (HSE), Communities and Local Government (CLG), Home Office/Ministry of Justice (HO/MoJ).

[2] In this report I have adopted the same definition of "science" as the Science Review programme, ie it covers both natural sciences and social sciences. Unless the context indicates otherwise, the references to "science" use this definition.

economists, operational researchers, social researchers and statisticians[3]. The focus of my review is such that the first two of its four recommendations are directed solely at GO-Science; but the second two are addressed – in whole or in part – to all of the professional disciplines. Many of the principles underpinning the proposed new model of science reviews could be applied to other disciplines.

[3] Each of these four professions has its own distinct service grouping within the civil service, eg Government Economic Service. There is no separate service for the science and engineering profession.

PART I: ASSESSMENT

6. This part of the report assesses the performance of the Science Review programme against three criteria.

(a) How well the programme has delivered against its overall aims.

(b) The extent to which the current Science Review approach and methodology has supported delivery of its aims.

(c) Whether it is cost-effective, that is the extent to which the benefits secured by the programme represent a good use of the resources it consumed.

(a) Delivery against overall aims

7. The overall aims of the programme were defined at its outset and have remained unchanged.

8. These aims reflect the objectives of the former Government Chief Scientific Adviser who judged that a detailed analytical and academic approach to scrutiny was the correct one, an approach endorsed by the Chief Scientific Advisers' Committee at the start of the programme. Although the two schemes have different goals and methods, there is a flavour of the Research Assessment Exercise for universities in the Science Review aims and process. It is no criticism of the original aims of the programme to say that the ways – including the styles - in which government business is conducted have changed in recent years, and so challenge the programme's own aims and ways of doing business.

First aim: Maintain and improve the quality and use of science in government

9. This is the central aim, and the most ambitious. If the Science Reviews are not contributing to its achievement, their whole rationale is in doubt.

10. The problem with an aim like this is measuring success. There is no shortage of opinions across Departments, Steering Panel members and stakeholders. At one end of this spectrum is the conviction that the experience of a Science Review had no impact on the way the Department did its business; at the other is the view that a Science Review put the use of science on senior management's agenda and improved the Department's performance in several respects, both strategic and day-to-day operational.

11. Yet there are no readily available hard indicators which enable clear links to be made between Science Review recommendations and changes in a Department's performance. The difficulties of developing such indicators are real, and would require a detailed understanding of the policy development processes in each Department in order to establish a worthwhile audit trail. It is beyond the scope of the Science Review programme to undertake such work, and this may account for the lack of penetration by Reviews into the ways in which commissioned science was actually used to influence the eventual policy decision[4].

[4] This issue is examined further in the Approach and Methodology section.

12. Other evidence is equivocal or absent. Departmental responses to Science Review recommendations frequently note that the action to implement a particular recommendation was already in hand as part of an internal business improvement project. The absence of follow-up reports from Departments detailing what has been done to implement agreed recommendations is a significant gap in the evidence base for an assessment of the programme's achievements.

13. It nonetheless remains possible to draw some tentative conclusions.

(a) Without the Science Review programme, and other measures taken by GO-Science and its predecessors such as the encouragement of Departmental Chief Scientific Adviser appointments, it is unlikely that measures to improve the quality and use of science in government would have been so consistently devised and advocated – and in many cases implemented.

(b) The great majority of recommendations made by the Science Reviews have been accepted by Departments, usually without qualification. The fact that the Reviews and Departmental responses are published and the existence of the convention that different parts of government do not argue with each other in public may have been an influence here. But it does suggest that the Reviews are providing helpful support to Departments' plans to improve the quality and use of science. There is evidence that the prospect of external scrutiny has led Departments to take action in advance of a Review to improve their compliance with good practices.

(c) The event of a Science Review puts science in the spotlight for at least some senior managers in a Department, though awareness of the Review at senior levels has been patchy. In one reviewed Department the Strategy Director had no knowledge of the Review.

(d) Some issues have recurred in successive reviews. Those which relate to weaknesses on particular projects – for example, an unsatisfactory commissioning process – may be failings by individual officials rather than a systemic problem. Others are clearly matters of corporate concern, for example the difficulties in recruiting or retaining scientific staff with the right skills. The latter provide evidence for GO-Science to follow up with Departments, including central Departments, as issues that might require concerted action if the quality and good use of science in government is to be maintained, let alone improved.

(e) The Science Review process can miss key risks in Departments. In one case the Department's Chief Scientific Adviser – who came into post in the closing stages of the Review – was disappointed that the Review had not picked up on what he and others in top management had themselves identified as the key problems affecting the Department's use of science. However since the Department's self-assessment at the start of the Review had not identified these problems, failure to engage effectively by the

Department is as much a contributory factor as any weaknesses in the Review process.

14. Achievement of the other overall aims of the programme is more susceptible to measurement.

Second aim: Review existing departmental systems for assuring the quality, management and use of their science.

15. Science Reviews examine Departmental systems and procedures on a range of topics with the aim of forming a view on the extent to which Departments meet the ten good practice criteria[5] that make up the common standards against which performance can be judged. The "peer review" element of the Reviews is the principal structured evidence source on the quality of the science, and the case studies in each Review provide more detailed information about the management of science, though there is some overlap between the two processes. Extensive information is also collected during interviews.

16. There is a tension between focussing on accepted Departmental practices and systems and on individual cases which is not always satisfactorily resolved in the Reviews. As an example, the Communities and Local Government Review examined the Department's tendering process, noted that competitive tendering was the norm, accepted that the case studies for the Review were "not typical of the Department's management of science", but nonetheless devoted substantial commentary to the procedural weaknesses in tendering for two projects. It is not clear from the report what wider lesson the Department was expected to draw from the management failings in the two case studies.

17. This example highlights a significant omission from the analyses in Reviews: the discussion of risk. Although things go wrong, organisations have to make judgements about the relationship between the consequences of a particular type of failure and the extent and cost of the measures to be put in place to prevent those failures. It would significantly add to the value of the Science Reviews if they offered some assistance to Departments to assess those risks or to manage them in a proportionate way.

18. Within the limitations inherent in examining a necessarily selective sample of projects out of a Department's total science activity, the Science Review programme has been able to make assessments of the quality and management of science that have been sufficiently soundly based to command the support of the Steering Panels who sign off the reports. Not all the conclusions are developed into recommendations for action: this is clearly sensible where the Review has found the Department is handling its science well, and it should enable the Review to reserve action points for issues considered to be of particular importance. Since Departments are asked to respond specifically to the recommendations, there is a risk that important conclusions not linked to a recommendation may be overlooked[6].

[5] See Annex D

[6] For example, the HSE Review commented that there was advantage in HSE consolidating some small projects into larger ones but makes no reference to this in the recommendations; and the HO/MoJ Review identifies the absence of succession planning to avoid loss of expertise and corporate memory as an issue for the Departments but makes no recommendation for addressing it.

19. The case studies which review a particular project in depth would be the obvious vehicle for examining how a Department makes use of science in identifying forward issues and in informing policy or delivery decisions. However, the extent to which the studies explore this part of the policy making cycle in any depth varies considerably. For example:

- some case studies provide some useful information about how the evidence was used to inform policy decisions or policy guidance, although an analysis of what other factors contributed to the final policy decisions would have enabled the reader to judge the relative weight being given to the science[7];
- other studies have highlighted failings by the Department to make sufficient use of the evidence[8];
- elsewhere, there is little or no discussion of how the science is used to inform decisions[9].

Third aim: Disseminate examples of good practice from within the UK and abroad.

20. The Science Reviews seek to identify examples of good practice within Departments and to highlight them in the pre-Review briefings, in the main reports and in discussions with individual officials during a Review. In some cases, the practice is explicitly commended to other Departments, although the mechanisms for ensuring that these commendations are picked up need further development. To date GO-Science has arranged two workshops, open to all Departments and focussing on the ten success criteria[10], and further workshops are in prospect.

21. That said, there is no stated process for deciding on whether an activity meets a "good practice" standard. Measurement against the ten criteria could be one such process, but this would not work where there is no direct relevance to any of the criteria. As an example, the Defra Review commended to other Departments the model of an overarching Science Advisory Council though without providing evidence of actual benefits delivered by the model.

22. The Science Reviews are not the only source of identifying good practice. The annual self-assessments by Departments against the "eight attributes" – see Annex D – generate examples of good practice which together with examples from the Science Reviews feed into GO-Science's dissemination activity.

23. GO-Science acknowledge that resource constraints have meant that work on disseminating good practice has not advanced as quickly or as effectively as they had hoped. In addition to the workshops, there are plans to collate the good practice examples on the GO-Science website, though the strong competition for the time of busy officials across government raises doubts about the effectiveness of this approach.

[7] For example, the case studies in the CLG Review
[8] For example, the Antimicrobial Resistance case study in the Defra review.
[9] For example, the Permissioning Regimes case study in the HSE Review.
[10] See Annex D

Achievement of the overall aims - summary

24. The Science Review programme is making some contribution towards the first of its overall objectives; and it is largely achieving that part of its second aim of reviewing Departmental assurance systems relating to the quality and management of science. During the period covered by my review, it did not effectively address the systems for the use of science by Departments, nor has it made much impact on disseminating good practice. Apart from the ever-present constraints on resources, the limitations on achievement stem principally from three factors.

(a) The first of the overall aims needs reformulating so that its achievement can be measured and that it bears a real-world relationship to what is deliverable through a scrutiny process. Scrutiny can help maintain and improve the quality and use of science in government but it is not – and should not be – the major driver.

(b) It is not clear – not only to me but also to many participants and stakeholders in the Science Review programme - for whose benefit the scrutiny is taking place. Is the primary intended beneficiary GO-Science (and the Government Chief Scientific Adviser) or the Department being reviewed? Or both in equal measure? I return to this issue in Part II.

(c) The methodology adopted by the Reviews is inhibiting exploration of some issues. I examine this in more detail in the next section.

(b) Approach and methodology

25. For the Departments that have already undergone a Science Review the experience has not been wholly satisfactory. In taking account of Departments' comments I have made allowance for their instinctive resistance to being scrutinised by the centre of government. Nonetheless, many of the Departments' concerns are echoed by others, including some Steering Panel members who approach the process objectively. And it is entirely proper to measure the approach of the Science Review programme against the expectations – set out by the top management of the civil service – of how the service should be operating in 2008.

26. The Science Review methodology has its strengths.

(a) As an external scrutiny mechanism, the Reviews provide a specific and structured opportunity for the Government Chief Scientific Adviser and senior Departmental management to form a view on a Department's use of science and to adopt improvements.

(b) In addition, middle management staff are able to raise issues that might not be acted on through internal line management channels.

(c) On topics where a Review does not bring new insights, it is nonetheless a valuable means of validating Departments' plans and actions, or of drawing attention to weaknesses or limitations.

(d) The Reviews are inclusive. The open consultation period means that anyone is able to make their views known to the Review team.

(e) The membership of the Steering Panels is impressive. Through them, the Reviews have the capability to apply serious and respected expertise and judgement to Departmental business.

(f) The staff work is thorough. Augmented by consultants if necessary, the Review teams assemble a formidable amount of evidence compiled from interviews, documents, and examination of specific projects.

27. There have also been some weaknesses. In order to understand how best to move forward, these need to be examined in some detail. They are categorised as follows.

(a) Engagement: Departments, Steering Panels and stakeholders.

(b) Proportionality: review criteria, peer reviews, case studies and interviews.

(c) Timing: scheduling and duration.

(d) Deliverability: recommendations, resources and traction.

28. The comments which follow are not to be taken as criticism of individuals. It is right to record here that several Departmental officials and Steering Panel members made a point of commenting favourably on the professionalism and expertise of the Science Review team.

Engagement

Departments

29. Although the formal processes of Science Reviews require engagement of the Department under review at Permanent Secretary level (the initial commissioning letter and the sign-off of the Department's response), the extent, nature and quality of Departmental engagement has varied significantly. For example:

(a) There has been little serious discussion of the Science Reviews at Departmental Board level. This suggests that the Reviews are not presenting themselves as an effective and relevant performance improvement facility.

(b) Departments attend Steering Panel meetings, but as observers not members. While their presence is clearly sensible for much of the time, for example to explain policies and practices, some Steering Panel members have found the current practice encourages Departments to be defensive and/or to exercise excessive influence at meetings.

30. Problems have arisen from an absence of shared endeavour. Departments which fail to engage seriously and constructively, and focus on defending their positions, will derive little benefit from the review process. Conversely, there is an obligation on GO-Science to demonstrate the explicit

benefits to the Department from the investment required in that degree of engagement.

Steering Panels

31. Steering Panel members are keen to make an effective contribution to the Reviews but some have raised concerns about their ability to do so.

(a) Although members have to date been prepared to sign off the final reports, there are mixed views about the extent to which they "own" the report. Some consider that the Panel has served primarily as a mechanism to validate the work of the Review team rather than actively to steer the Review. Others note that given the way in which the Reviews are structured, this is a reasonable approach.

(b) The length of time between Panel meetings means that members have to re-learn the Review issues and material. This is exacerbated when papers – which can be lengthy – are circulated only a few days ahead of a meeting.

(c) As a rule, there has been no serious opportunity for Panel members to meet collectively as a team for open and frank discussion, without Departmental representatives present.

Stakeholders

32. As noted above, the Reviews adopt an inclusive approach to engaging stakeholders. While this is a strength and is clearly welcomed by those stakeholders who take the opportunity to contribute, there are risks in the current process and little evidence that these risks are effectively managed.

(a) The consultation process is undertaken at an early stage in a Review, typically open for 12 weeks. The results are usually reported to the first Steering Panel meeting. Consequently, issues raised in consultation can strongly influence both the initial fieldwork and the proposals for the final report before the Panel has had an opportunity to give a steer.

(b) Departments have expressed concern, sometimes forcefully, that there is no critical evaluation or weighting applied to stakeholder comments.

(c) The response rate[11] has ranged from 6% to 23%. This is not out of line with many other consultation exercises, but given the importance of the consultation in the Review process it increases the risk of bias towards those who want to make criticisms as opposed to those who see no need to comment because they consider the Department is performing well.

[11] Figures from the Defra, HSE, CLG and HO/MoJ reviews: the DCMS review adopted a different – and wholly unsuccessful approach – to general consultation.

Proportionality

33. With modern ICT the opportunities for collecting and assembling information have become vast. To keep any form of review procedure manageable, judgements need to be made not only about what sort of information should be used to inform the conclusions but also how much of it is needed.

Review criteria

34. From their inception, Science Reviews have sought to bring consistency between reviews by applying a set of ten criteria[12] which would give indications about the quality and use of science in Departments. The original intention was that a particular Review would focus on a selection of the criteria, but in practice Steering Panels and stakeholders have argued that all ten were important and that the Review should address them all.

35. This has introduced a degree of rigidity into the Review format and has made it difficult to comply with GO-Science's stated policy that a "one-size-fits-all" model is not an objective of the programme.

Case studies

36. Reviews typically select four areas for detailed examination, with individual Steering Panel members exercising some oversight of the work. Earlier Reviews engaged consultancy firms for this work, though the results were mixed and GO-Science has since made less use of consultants.

37. The case studies are resource-intensive, for GO-Science, Panel members and the Department. The stated justification for them is that they validate – or otherwise – the findings from other sources (eg interviews, documents). They are written up in considerable detail, often including several pages detailing the policy objectives and other background which are not directly relevant to the findings.

38. The selection of case studies is a major issue for the initial stages of a Review, with considerable efforts made to find topics which satisfy certain standard criteria agreed with Departments. Although these criteria are offered for discussion by the Steering Panel – who might suggest additional criteria - some Panel members see them as a further constraint on their ability to steer the Review. There is general acceptance that the case studies are unlikely to be fully representative of a Department's business, particularly where the Department engages with a wide range of science.

39. Analysis of the case studies – see Annex C - shows that the extent to which the case studies inform the recommendations in the report is variable. Both the case studies and the peer reviews generated findings in relation to criteria 4 (commissioning and managing new research) and 5 (ensuring quality and relevance of science) of the ten criteria, but relatively few on the others and none at all for some. The main report of the Home Office/Ministry of Justice Review was highly critical of the way in which science was used by policy-

[12] See Annex D

makers, yet two[13] of the three case studies found that science was being used well to inform policy. All this raises questions about whether the case study approach – with the limitations inherent in a small sample – is adding sufficient value in its present form.

Peer reviews

40. The expression "peer review" as applied to the process used in the Science Reviews is a misnomer and has caused some confusion with the peer review of projects used at either the commissioning or evaluation stages. It has led Departments to comment that since their projects have already been peer reviewed, there is no point in doing so again.

41. The peer review stage of the Science Reviews aims to look at the quality and fitness for purpose of science commissioned and used by Departments by applying a set of standard questions to about 10 science projects. The standard questions have evolved from Review to Review following advice from each Steering Panel. There has been some comment from Steering Panel members that this is essentially a tick-box process, with the concomitant rigidities, and that better results would be achieved by allowing greater scope for comments.

42. The analysis of the peer reviews reported in Annex C enables similar conclusions to those for the case studies (above).

Interviews

43. Information obtained at interview forms a significant part of the evidence base for Reviews[14]. Interviews are typically conducted by two members of the Review team, covering the criteria relevant to the interviewee in a methodical way, and are written up in detail (the writing up can take anything between half a day and a full day). The stated justification for this is to enable any finding or comment to be sourced accurately should it be challenged.

44. This contrasts, for example, with the interview style of Office of Government Commerce Gateway reviews[15]. Here, typically, interviews are conducted by one, two or all three members of the team, who make rough notes to inform the evidence base for the report. Initial questions are often standardised, to identify issues for more detailed exploration. The draft report – provided to the SRO[16] for comment before it is finalised – is normally the only formal record. The interview notes themselves are not formally written up, and there is no evidence that this causes problems in substantiating statements in what can be very critical reports.

[13] NOMS and North West Offender Management Pathfinder case study; Prolific and Other Priority Offender (PPO) Programme case study

[14] Figures are in Annex C.

[15] A "peer review" process in which independent practitioners from outside a programme or project use their experience and expertise to examine the progress and likelihood of successful delivery. Reviews are used to provide an additional perspective on the issues facing the internal team, and an external challenge to the robustness of plans and processes. OGC Gateway reviews are best practice in central civil government, the health sector, local government and defence.

[16] Senior Responsible Owner, the person responsible for ensuring that a programme or project meets its objectives and delivers the projected benefits.

Timing

Duration

45. There is general agreement that the Reviews have been too long drawn out. The original design and application of the Review process is now out of step with today's civil service imperatives of "pace and urgency". Reviews have typically taken some 18 months to complete, on occasions up to 2 years. This is not always a result of Review team decisions: there are examples of Departments delaying access to their staff and of engaging in lengthy arguments about the drafting of the report. However the Review structure used to date is inherently lengthy, with 150 days allocated for 4 case studies and 10 peer reviews alone.

46. The Review team is seeking to speed up reviews, but work carried out in 2006 concluded that, without an increase in staffing resources in the team, there is an irreducible minimum of time required by the present format and content. More recently, proposals were made to limit the Food Standards Agency review to 8 months, but following discussions between GO-Science and the Agency it was agreed to extend the review to 11 months.

47. Within government, comparisons can be made.

(a) The first round of Capability Reviews took three months per Department.

(b) Defra aims to complete each five-yearly science audits of its laboratory agencies in about two weeks, though with teams averaging about 20 independent experts.

(c) OGC Gateway Reviews typically review a major high-risk project or programme in one week (including preparation time): although the focus is different, the Gateway process shows what can be achieved in a short time even when the consequences of a review getting it wrong are serious.

(d) Audits carried out by the UK Accreditation Service to check compliance by research contractors with the Defra/Food Standards Agency Joint Code of Practice for Research involve typically a single day on site, plus preparation and report-writing time. Again, the purposes are different, but the UKAS method shows that a single project or set of projects can be reviewed for compliance with agreed standards in a short time.

48. It is sometimes argued that the long lead times involved in some scientific projects - whether the duration of the project itself or the time elapsed before the impacts of the work can be seen – mean that a short review process is unsuitable. This is not a compelling argument: if a project is not complete or its impact is uncertain, then other projects should be selected for review. The volume of science commissioned and used by Departments provides an ample pool.

Scheduling

49. "Review fatigue" is not new in the public sector, but Departments (including the centre) and external observers have commented to me that the pressures are stacking up. For example, in the past two years the Department of Health's strategy and science teams have been on the receiving end of, successively, the Cooksey Review[17], a Capability Review[18], and a Science Review[19]

50. GO-Science seeks to schedule Reviews taking account of several factors, including: alternating large and small Departments, issues of particular significance current in a Department, whether the GCSA has any particular concerns about a Department, and the availability of Departmental resources to support the Review.

Deliverability

51. The operational focus of a completed Science Review is inevitably on the recommendations. Those are the action points, which Departments are expected to implement (or explain why they are not to be implemented). The impact of those recommendations is therefore a key success measure for the Science Review programme.

52. It follows that making recommendations which are deliverable in practice is essential to the credibility of a Review. This is now a standard requirement for policy and regulatory proposals, and the Science Reviews could usefully adopt a similar discipline.

Recommendations and resources

53. There is a tension between what is desirable and what is deliverable. It is a function of a review process to point towards the former but this needs to be done against a recognition of the real business environment. Getting that balance right is difficult but important. The Science Reviews have not always succeeded, in two particular respects.

(a) Some recommendations are too vague to be helpful, or are no more than a statement of common sense[20]. They send the message that the Review has found a problem but has no solution in mind. Implementation cannot be measured.

(b) The resource consequences of recommendations are normally ignored[21]. This is unhelpful when running cost budgets are under

[17] Sir David Cooksey's review of UK health research funding, commissioned by the former Chancellor of the Exchequer, which ran from April to December 2006.
[18] First half of 2007.
[19] Commenced July 2007.
[20] For example: "The Department should identify effective ways of ensuring that the principles of the Guidelines are embedded in the processes for managing and using its analytical work"; "There is a need for better knowledge transfer between professions, especially between the natural and social sciences and between scientists and policy makers."
[21] For example: "Secondments of social scientists from OGDs (e.g. from CLG) and academia might also help in the migration area"; "As a starting point for systematically sharing knowledge, CLG should consider implementing an internal online management information system for all analytical work and research that is commissioned".

extreme pressure (as they have been in most Departments during the life of the programme). It is argued, with justification, that GO-Science cannot decide on a Department's spending priorities and that it is for Departments to decide what priority to give to Review recommendations with resource consequences. At the same time, some recognition of competing priorities, some indication of the range of costs involved and some assessment of the risks of not implementing the recommendation would greatly enhance the helpfulness and credibility of the Review.

Implementation levers

54. Recommendations are made to be implemented. As noted above, engaging Departments with Science Reviews at top management level has been problematic. Although it will often be possible to implement second-order process recommendations without the need for Board approval, significant changes – particularly those with resource consequences – will need to become part of the Department's high-level change agenda if they are to be implemented effectively.

55. GO-Science is aware of Departmental concerns that some Reviews contain too many recommendations addressed to the Department as a whole, ranging from strategic to operational detail. As a response, the Home Office/Ministry of Justice Review included a table showing the intended recipients of each recommendation, and the forthcoming report of the Department of Health Review plans to include a similar facility. Departmental reaction to the HO/MoJ table has been mixed.

56. Some recommendations recognise the issue and include an explicit link with a key Departmental business process, such as the business plan or the evidence and innovation strategy. Others take their chances, relying on a sympathetic sponsor to push them forward. For example, the Home Office/Ministry of Justice review addressed six recommendations explicitly to the Permanent Secretary, but only one was structured for leverage by linking to a corporate business process, in this case the Departmental business plan.

57. Securing this sort of traction is a further factor in the scheduling of future Reviews where implementation is dependent on other business process cycles.

Effectiveness of approach and methodology - summary

58. The Science Review approach has scored some significant achievements which should not be undervalued. Yet it has weaknesses which have adversely affected the programme's reputation. Some of these are recognised, and GO-Science has sought to make improvements. However, I conclude from my discussions with Departments, steering panel members and stakeholders that these will be insufficient to restore the programme's credibility and so maximise its impact. A relaunch of a redesigned programme is the only realistic way of building on the achievements to date and of speeding up the individual reviews, and I make recommendations on this in Part II below.

(c) Cost-effectiveness

59. The costs to government of running the Science Review programme are not accurately quantifiable because of the difficulties in estimating the staff time and other costs in the Departments reviewed, though these are likely to be in the range of £15k-£20k per review. The costs to GO-Science over the past 3 years are:

	2005-06	2006-07	2007-08
Admin	472	486	607
Programme	477	211	125
TOTAL	949	697	732
Figures are £000. Admin spending covers staffing, staff-related and central services costs. Programme spending covers consultancy, research, and Panel and peer review expenses.			

The reduction in programme costs from April 2006 is largely due to the decision to carry out more of the fieldwork in-house rather than contract it out to consultancy firms.

60. My review has not undertaken a formal cost-benefit analysis. This would yield results of little value, both because of uncertainties around Departmental costs and because of the difficulties in separating out benefits from changes driven by a Science Review from those which Departments intended to make in any case

61. The judgement to be made is whether the expenditure to date has been value for money. Given the ambition of the programme, the costs are not unreasonable. Yet the shortcomings in realising that ambition, identified above, mean that greater benefit could be secured from the expenditure if changes to the programme are put in place.

PART II: MOVING FORWARD

62. This part of my report sets out an agenda for what should be seen as the next generation of reviews of the use of science, and indeed of evidence more generally, by Departments in delivering their objectives. It proposes a mix of incremental changes and clean breaks from the past. The agenda builds on the achievements of what has gone before while addressing the factors that have prevented the Science Reviews from achieving their full potential in supporting performance improvement across government. As noted above, GO-Science has started to introduce changes to the review process, being tailored to and agreed with the Department being reviewed. Although these changes are in the right direction, a more radical programme of change would deliver significantly greater benefits.

Principles

63. The proposals are built on five principles:

(a) Clarity of purpose. Reviews should be designed and executed with a clear and defined objective, which may vary between reviews. Where there are multiple objectives, there should be no explicit or implicit incompatibility between them.

(b) Sharpness of focus. Reviews should only undertake activity which directly delivers the review objectives, and be demonstrably managed economically and efficiently.

(c) Flexibility. A sharp focus does not mean rigidity. Reviews need to be able to take into account the differing activities and priorities of Departments and be structured accordingly.

(d) Integrated delivery. Reviews should be aligned, in content and timing, with other quality assurance and performance improvement mechanisms.

(e) Capacity to evolve. Reviews should develop in form and content, as the needs of government adjust to perpetually changing circumstances.

(a) Clarity of Purpose

64. There is no common perception about whether the Reviews are a form of audit or inspection, a critical friend, a developmental tool, or a combination of all three. This has led to confusion, delay and unrealised expectations on all sides.

65. The first key to progress is to resolve the tension between, on the one hand, the needs of the Government Chief Scientific Adviser (GCSA) to be able to reassure the Prime Minister about the "health" of government-funded science and, on the other, the expectations of Departments that a scrutiny process should deliver benefits to them at least commensurate with the resources consumed by the process.

66. The second key to progress is to decide whether the purpose of the reviews is:

- to assess a Department's capability in the commissioning and deployment of science, or
- to appraise the way in which science has been used in policy development, or
- a mix of these.

67. Resolution of these issues centres on the respective responsibilities of the GCSA[22] and Departmental Permanent Secretaries[23]. It is the Permanent Secretary's role to ensure a Department puts in place systems that are fit for the purposes of securing relevant scientific evidence of an acceptable quality and of ensuring that this evidence is used to support forward planning and decision-making, and that it provides the resources to operate them effectively. It is the GCSA's role to assess the extent to which the culture of a Department and the systems it puts in place help the achievement of these objectives, and to offer support where Departments need it.

68. In practice, these accountabilities represent different strands of an integral set of processes – good planning, stewardship and decision-making. In ensuring that the right science is available and is used by a Department, the GCSA and the Departmental Permanent Secretary have a shared objective. It follows that they have a shared interest in ensuring that the systems to achieve this are put in place and that the culture exists to make them work effectively. Operating in partnership does not, and should not, remove the element of challenge when scrutinising performance.

69. There is by now a substantial body of guidance and good practice, some of it generated by the Science Review programme, against which the GCSA can judge the fitness of Departmental systems generally and Permanent Secretaries can do the same for their own Departments. There has been significant change across government in attitudes to the importance of good science, and of good evidence generally, since the cross-cutting review of science and research reported in 2002 when public confidence in government's use of science was at a low ebb following, in particular, the BSE and foot-and-mouth outbreaks and enquiries.

70. The appointment of Chief Scientific Advisers in Departments, championed by GO-Science and its predecessors, has been a particularly significant driver. The CSAs have brought a strong academic focus to organisations which are in the nature often driven by short-term imperatives. The changes put in place include the wider use of better systems for identifying the need for science, for assuring its quality and relevance, and for applying it to policy, regulatory and delivery decisions. Progress has been made, however patchy, in aligning science strategies with Departmental priorities and business plans.

[22] And as necessary the Government Chief Social Researcher. Further references to the GCSA should be taken to include the GCSR as appropriate.

[23] Here, and elsewhere, references to Departmental Permanent Secretaries reflect her or his position as head of the Department and carrying the formal responsibilities of that post. In practice, as is normal, the Permanent Secretary will take advice from senior colleagues: in this context from the Department's Chief Scientific Adviser, heads of profession, and chief analyst, according to the Department's organisation; and advice on decisions will be prepared by staff work between GO-Science and the Department.

71. Against this background, it is fair to question whether the objectives of the central scrutiny put in place following the cross-cutting review are still the right ones. Part I of this report has identified weaknesses in current processes, some of which stem from a lack of clarity about the purposes of the scrutiny as the Whitehall world moves on. It also noted that achievement of the first of the current overall aims is very difficult to measure. GO-Science are in the process of reviewing the analytical framework for monitoring the health and use of science in government, and their current thinking is reflected in my proposals.

Recommendation 1: The overall objectives of the GO-Science review programme should be redefined as follows:

To work with departments:

- to check that they have in place the capability, systems and culture to access and use science well to support strategy and policy development, delivery and risk management; and
- to provide targeted expert support to improve in these areas.

72. How these objectives can be delivered is outlined in the following paragraphs. What is important is that the review programme is seen as an integral part of GO-Science's wider efforts to improve the access and use of science by government, and indeed of the ambition across all the professions to drive further the shift to an evidence-based culture. The review process must become – and should be presented as such – part of a continuing dynamic to address evidence and analysis issues within and across Departments. This implies regular contact with Departments between reviews, but in a supportive rather than progress-chasing mode.

(b) Sharpness of focus

73. The new objectives proposed above, with the emphasis on checking the culture, capability and systems rather than in-depth examination of individual cases, open the way for a more focussed form of review, of much shorter duration and reduced cost. The savings would enable more reviews per year to be carried out, and so provide more regular and up-to-date assurance than is possible at present.

74. There are two distinct elements to the new objective which should be separately identified in any review project plan to ensure that focus on both is not lost. Those elements are *accessing* science and *using* science. The first will focus on how well a Department plans its science base and then populates it with evidence of acceptable relevance and quality. The second will focus on how a Department uses science to discharge its functions, ideally probing such issues as the relative weight given to science – and "hard" evidence generally – in comparison to political or financial factors.

75. To some extent these are two discrete processes, but because there is a feedback loop through which what is done with the science should inform the business of securing science of the right quality and relevance in the future, there is a strong case for assessing both processes within a single review

framework. It involves Departments being willing to open up their decision-making processes to external review and, particularly where advice to Ministers is concerned, there may be constraints on what can be published in the report of a review. This does not lessen the value of a well-planned and executed external scrutiny, since sensitive detail could be reported to the Department privately.

76. Though shorter, this would not simply be a tick-box review. Structured (but not standardised - see section (c) below) approaches would be appropriate for much of it. Both to secure credibility and to provide a high quality review, outside expert practitioners would continue to be engaged, though in a hands-on role akin to external reviewers on Capability Review, Gateway or research unit visitation teams, with adequate staff support from GO-Science. The hands-on involvement of the external experts, bringing both breadth and depth to their engagement with the Department, would raise significantly the learning dynamic in the reviews. There would be no Steering Panel as in the present model. An outline of what this type of review might look like is at Annex A.

77. A review process to which there is no follow-up lacks teeth. There needs to be a mechanism for responding to the findings of those new short focussed science reviews which reveal weaknesses in Departments. Hence the second part of the proposed new overall objective for the review programme. To achieve this objective, Departments with revealed weaknesses would be expected to accept a more thorough review focussing on those areas where problems had been identified, though not necessarily confined to them. This type of review would be no less focussed than the shorter reviews but it would offer Departments the opportunity for other relevant issues to be examined as part of the review, with the agreement of GO-Science. External experts would again be involved. Because the review would be in response to weaknesses identified in a Department and would be providing expert support to help address them, it is reasonable that the Department should contribute to the costs. A brief outline of the model is included in Annex A.

78. The proposal for two types of science review is not a menu of choices. The Office of National Statistics has used a review process based on three categories of review, which they found to be not wholly successful. It is an important element of my overall package of recommendations that the short and detailed scrutiny reviews should be used sequentially, other than in exceptional cases. Jumping straight to a detailed scrutiny review without having clearly established the need for one through a short review risks reinventing the existing form of Science Review as well as wasting public money.

(c) Flexibility

79. Rigidity has big risks: issues are missed, unnecessary activity takes place, costs rise. Rather than, as now, be constrained by a set of fixed criteria and another set of largely fixed process steps, the nature and shape of both the short and detailed reviews would vary. Factors influencing the design would include the size of the Department's science activity, its importance to achieving Departmental objectives, areas of particular concern known in advance to the Department or GO-Science, and the views of the external experts.

80. This is particularly important in securing traction for a review's recommended actions. The closer these recommendations can be aligned to

other activities to which the Department is committed – and the business plan and current strategy documents will be key sources of information – the greater the likelihood that science issues will be taken into the mainstream of Departmental programmes.

81. The current Review programme has focussed on Departments rather than on their agencies and other arm's length bodies. The process of shifting responsibility, funding and therefore risk to arm's length bodies has accelerated in recent years. The need for scrutiny of the acquisition and use of science by arm's length bodies will vary significantly, but it would be sensible for GO-Science and Departments to start considering now on a case by case basis whether what I am proposing should be extended to them.

82. Making changes of the sort proposed carries some risk. A short review invites the criticism that if an 18 month review cannot provide a wholly satisfactory account of a Department, what chance is there in three months? The answer is that there is no certainty – now and under the new proposals – that a scrutiny will pick up everything. What my proposals seek to do is rectify the weaknesses in the current system and secure better value for the public expenditure applied to it. It will be essential to monitor the new approach and be ready to make further changes in the light of experience.

Recommendation 2: A two-tier scrutiny process should replace the current standardised model for the GO-Science reviews. A ***short review*** focussed on Departmental practices and with a major role for external experts would be carried out in all Departments, with the option of extending them to arm's length bodies if the sponsor Department were to contribute to the cost. Significant weaknesses would be the subject of a more ***detailed scrutiny review***, for which a Departmental contribution to costs would normally be expected. Both types of review would be managed by GO-Science on behalf of both the GCSA and the Department's Permanent Secretary.

(d) Integrated delivery

83. Ensuring Departments are not asked to participate in a series of separate reviews covering similar or linked territory should be a basic function of the centre of government.

84. At present, Capability Reviews are the performance improvement vehicle with the greatest impact on Departments, primarily because of their sponsorship by the Cabinet Secretary personally. The programme of initial reviews was completed by December 2007; and a programme of re-reviews is now under way for completion by July 2009. There is a widespread view that in the first round the findings of the Capability Reviews on the "Base choices on evidence" element of the capability model (S2) were not particularly strong or convincing.

85. As noted in Part I, well-founded recommendations from science reviews are likely to achieve more traction if they can be linked to a high priority Departmental business process. The new short science reviews proposed in recommendation 2 offer a time-limited opportunity to make those links through the Capability Re-Reviews.

86. Running a short science review in the run-up to a Capability Re-Review offers advantages to the Department , the Cabinet Office and GO-Science.

(a) The Department benefits by having review activity brigaded within a narrower window.

(b) Cabinet Office benefits by having the relevant findings of the short science review available to inform the Re-Review.

(c) GO-Science benefits by having the opportunity of issues for action on the use of science by the reviewed Department included in the Capability Re-Review action agenda.

87. That said, it would be unwise to make too much of this. First, significant weaknesses in a Department's science systems would need to be examined in more detail by the proposed second-tier of science review, and in any case such detailed examination would fall outside the scope of the Capability Re-review. Second, it is not known what will happen after July 2009 when the Capability Re-Review round is complete; and so linking the new science review process to the Capability Review process is at best a short-term expedient. Indeed, the time needed to put in place new arrangements within GO-Science to support the new types of review could mean quite simply that the boat will have sailed.

88. The principle of securing better alignment with other reviews remains important. The geometry of relationships between science reviews and the other professional assurance systems for statisticians, social researchers, economists and operational researchers is variable. A summary of these processes is at Annex E. Irrespective of possible developments across the analytical professions (see recommendation 4), it makes sense to attempt to coordinate the timing of scrutinies of Departments' evidence and analysis functions. The benefits outlined at paragraph 86 would read-across to this sort of co-ordination.

Recommendation 3: Both as a short-term benefit and as an experiment in aligning different reviews for mutual support:

- existing plans for reviews of Departments by the analytical professions should be mapped and scheduled wherever possible to minimise the load on the Department and enable the reviews to benefit from the findings of others (Heads of Analysis are well-placed to agree who should lead on this work);
- GO-Science should consider the scope for running one of the new short reviews ahead of the start of any suitable Capability Re-Reviews from late 2008 onwards.

GO-Science should intensify their engagement with Cabinet Office to influence their thinking about the performance review landscape after Capability Reviews so that the potential of the new science reviews can be fully exploited, evidence and analysis can be properly integrated into any central performance review mechanism, and plans for further desirable modifications to the programme can be developed.

(e) Capacity to evolve

89. The proposed two-tier model of science reviews is not intended to be set in stone. As government evolves, so scrutiny mechanisms need to keep pace.

90. A significantly accelerating trend is joint working across Departmental boundaries. This is not new, but the PSA delivery plans following CSR 2007 introduce a greater requirement than hitherto for joint working by Departments. This inevitably raises questions about how well-equipped Departments are to work jointly on the commissioning and use of science and other evidence, particularly where different traditions and practices exist in the partner Departments.

91. The Science Reviews have identified some recurrent issues in Departments which suggest that further developmental activity is needed. These include staff recruitment and retention, horizon scanning, communications and knowledge management. There is some support for supplementing Departmentally-focussed scrutiny with occasional thematic reviews looking at how such issues are being managed and providing a vehicle for targeted effort to resolve outstanding problems with them. These reviews need not be limited to science: where the issue impacts on other disciplines, then a multi-disciplinary review would add value.

92. In parallel, many Departments have brought, or are bringing, together their professional staff of different disciplines who handle evidence and analysis into combined multi-disciplinary units. This development recognises the importance of identifying all relevant evidence sources, actual and potential, and of framing policy questions on a multi-disciplinary basis. It is expected to lead to the development of systems for commissioning and managing evidence which are common across the disciplines.

93. The assurance systems for the different professions have different purposes, and so there are no immediate opportunities for integrating them. Indeed there would be difficulties in doing so in relation to the assessment of professional competence and quality, since a scientist could not sensibly sit in judgement on a piece of economic analysis.

94. However, the new two-tier model of science reviews will generate fresh expertise in reviewing Departmental systems. Relevant parts of that expertise would be transferable to other disciplines, just as some of the assurance processes used by the other disciplines are transferable to science.

95. Provided progress is not mired by professional rivalries, there is an opportunity to develop the new science reviews and the assurance processes in other disciplines into a model which operates in new cross-cutting dimensions.

Recommendation 4: Heads of Analysis should sponsor, on a pilot basis, a programme of new cross-cutting assurance reviews which may include:

(a) A model which helps Departments develop systems for joint working on evidence and analysis on topics where there are joint responsibilities for delivery.

(b) A model which enables attention to be focussed on "evidence management" issues common to several Departments.

(c) A model which enables relevant Departmental issues and practices to be scrutinised by a single review covering all the disciplines.

(d) A model which combines both (a) and (c) or (b) and (c).

96. Further consideration will be needed about the ownership of such reviews. The model of GCSA and Departmental Permanent Secretary may not be the optimum in all cases, for example where the review covers work directly related to a joint PSA.

Delivery, traction and sustaining improvement

97. One of the underlying messages of my review is that the business of scrutinising science is central to the business of government and that it deserves a review mechanism of the quality and status commensurate with its importance. The current mechanism is, at best, regarded as helpful but tangential to the "real" business.

98. An assurance system that is respected and acted on by Departments is not only the GCSA's ambition: it is widely shared around Whitehall. Adoption of the recommendations in this report is the next step to developing such a system. The drive for a focussed but tailored review model, with greater use of external experts to enhance credibility, authority and value, will help kick-start the virtuous circle of making science reviewing an attractive and exciting business for civil servants to be in.

99. In particular, by making key linkages with other business drivers wherever possible – and GO-Science should be opportunistic about this to the extent that resources allow – the traction for ratcheting up performance in accessing and using science will develop. The partnership basis for reviews which gives Departments and GO-Science an equal stake is the bedrock of this change, because it will make the links to the Department's priorities and business plans. Alignment with other assurance processes not only reduces demands on Departments, but also opens up opportunities for greater working across professional disciplines and so increases the efficiency and the impact of the assurance business as a whole.

100. The importance of planning the move to the new model and of ensuring that its roll-out and subsequent performance is properly monitored cannot be over-emphasised. As a first step, new governance arrangements should replace the current GO-Science programme board (which is currently constituted as an internal management meeting). What these arrangements should be needs further consideration, but they should at least:

(a) be consistent with OGC guidance on programme management; and

(b) provide for direct input into the development, management and monitoring of the new review programme by other Departments and from outside of the civil service.

101. If my recommendations are accepted, there is a substantial implementation programme to be developed and executed. In addition to the specific points identified in this report, implementation issues will include:

(a) Fuller development of the review model, including determining the frequency of short reviews, iteratively with (b) below.

(b) Alignment between the skills and resources needed to deliver the review model and the skills and resources – both from within the civil service and outside - that can be made available.

(c) Development of a forward programme of reviews.

(d) A new name for the GO-Science review programme, and its launch.

(e) Identifying adequate staff support to Heads of Analysis to deliver Recommendation 4 and their interest in Recommendation 3.

ACKNOWLEDGEMENTS

102. I am very grateful to the many people, inside and outside of government, who made the time to share their views with me and so enable me to draw conclusions informed by practical experience. In particular, the differing perspectives brought by members of the review steering group have been invaluable in teasing out the underlying – and not always wholly reconcilable – tensions.

103. I want to identify one individual: Dr Patrick Erwin, the outgoing head of the Science Review Team at GO-Science. His formal duties have been to serve as secretary to the steering group and to project manage the review inside government. But, less formally, he has acted variously as information provider, sounding board and wise counsellor. In these roles, coupled with his readiness to see the opportunities presented by change, he has provided invaluable support to the review.

104. The report and the recommendations are of course my own responsibility.

Peter Cleasby

ANNEX A

OUTLINE OF THE NEW REVIEW MODEL

1. This note sets out principles and characteristics of the new review model. As proposed in the main report, the new model should be sufficiently flexible to respond to differing circumstances. It follows that a single detailed blueprint will be unnecessary, although the model is likely to need further elaboration during the implementation phase of the changes. The planning stage for each review will be the vehicle for settling the detailed design on each occasion.

2. The new review model has two elements: the short review and – if it is required - the detailed scrutiny review. The objectives and characteristics of the new model are not confined to science reviews, and can be applied to reviews in other professions and to multi-disciplinary or other cross-cutting reviews[24]. The proposed pilots for cross-cutting reviews (Recommendation 4) would normally be short reviews.

Objectives

3. The review process would deliver the new objective of the review programme (To work with departments to check that they have in place the capability, systems and culture to access and use science well to support strategy and policy development, delivery and risk management; and to provide targeted expert support to improve in these areas)[25]. In particular, it would establish how far Departments are:

   (a) taking a long-term, strategic approach that takes account of science and research (both within the department and across Government);

   (b) commissioning, managing, quality assuring and evaluating science and research well;

   (c) ensuring that policy, regulatory and other decisions are based on good quality scientific evidence and advice;

   (d) managing and communicating knowledge effectively;

   (e) fostering and maintaining their internal and external scientific capacity.

4. These would be the objectives of both the short review and, in more depth, of the detailed scrutiny review where one takes place.

Design principles

5. Each of the short and detailed scrutiny reviews would be designed in line with the first three of the five principles set out in Part II of the main report: clarity of purpose, sharpness of focus, and flexibility. The other two principles - integrated delivery, and capacity to evolve – will be applied to the design of the overall programme of reviews in each business planning cycle.

[24] For simplicity, this annex describes the model in terms of a science-focussed review.
[25] Recommendation 1 of this report

Approach and characteristics

Short reviews

6. Unless the circumstances are exceptional, the following features should be present in the review:

   (a) At the outset, the specific objectives and terms of reference for the review will be agreed between the GCSA[26] and the Departmental Permanent Secretary (on advice from the Department's Chief Scientific Adviser and the Chief Analyst). These will take account of:

      (i) issues already identified by GO-Science (eg from the Department's self-assessment, or specific evidence management topics causing concern in one or more Departments); and

      (ii) issues on which the Department is seeking independent scrutiny and support (eg, on evidence-related issues identified in the business plan, or arising from well-informed stakeholder comments); and

      (iii) the availability of resources (to ensure GO-Science resources are managed efficiently, there should be a standard indicative budget for each review – which will in part be related to the duration – which should not be exceeded other than in exceptional circumstances).

   (b) The external experts will be appointed by the GCSA, with the Department's agreement. Their selection should be informed by the specific issues to be addressed in the review. The number of experts will normally be between 3 and 5, though a review with a broad scope – for example of a large and particularly diverse Department – may require more. Each team would include a serving Departmental CSA. To reduce "learning curve" difficulties, GO-Science should consider establishing a pool of external experts who would become familiar with the standard elements of short review methodology: senior academics recently retired from full-time work could be a useful source of such people. The pool might provide up to 2 of the experts, with the rest being drawn from relevant specialisms as necessary. At least half of the experts in a review will need to be able to probe the way in which science has been used to inform policies, as opposed to assessing the ways in which the Department obtains science of the necessary quality. A brief summary of key skills and attributes for all external experts is at Annex B.

   (c) The GO-Science team will be assembled in parallel.

   (d) The Department, GO-Science and the external experts should agree the review plan, including the indicators to be used to measure the extent to which the Department is meeting the criteria listed under

[26] And as necessary the Government Chief Social Researcher. Further references to the GCSA should be taken to include the GCSR as appropriate.

"objective" above. The expectation is that GO-Science will develop a standard set of indicators, making use of measurable metrics as far as possible (and there may be lessons to learned from the recently announced changes to replace the higher education Research Assessment Exercise[27] with a more metrics-based approach). The plan should be developed so that no more than 10 weeks are allowed from the appointment of the experts to the production of the report. It will need to ensure clarity between those parts of the review looking at how science is accessed and how it is used.

(e) The preparatory stage of the review will involve the GO-Science review team – under guidance from the external experts - identifying relevant documentation, preparing initial assessments of compliance with systems and performance against metrics and mandated or advised procedures, and arranging the programme of interviews between officials/a small number of key stakeholders and external experts. There should be no difficulties in obtaining necessary information because the team will be working to a plan agreed by the Department.

(f) The assessment stage of the review will involve the external experts, briefed by the GO-Science team on the findings to date, spending up to 10 days in the Department interviewing officials – who will normally be senior – as well as key partners and stakeholders, and examining key written evidence. In some cases, short telephone interviews will be both necessary and sufficient. The external experts will be forming judgements about the culture of the Department and its attitude to using evidence, as well as about the quality of the evidence base. They will need to agree in each case how to divide up the work between them. The aim will be to collect sufficient evidence for the experts to reach conclusions after applying their expert judgement: there will be no time to "gold-plate" the evidence base.

(g) The review report will be drafted by the review team, cleared by the external experts and submitted to GCSA and Departmental Permanent Secretary for sign-off. The report should be

(i) brief, limited to a summary of how well the Department measures up against the expected standards. Recommendations should focus on major issues and be limited to those requiring senior management action. Advice on minor processes or good practice should be included in a side annex, without the status of formal recommendations.

(ii) agreed by the external experts, the GCSA and the Departmental Permanent Secretary, though drafting arguments on points of detail should be avoided. If the Permanent Secretary and the GCSA cannot agree on a recommendation, then it will be for the

[27] See DIUS news release of 24 April 2008, *Denham Announces Next Steps to Assess Research Excellence.*

GCSA to make the recommendation and for the Department to explain why it disagrees.

(iii) made public, though there is no obvious need for the glossy format of the current Science Review reports to be continued. A clearly presented pdf document on relevant websites should suffice.

(h) Where recommendations have been made, the Departmental response and action plan should be made public within 4 weeks of the report being submitted. Again, the expectation should be for this be agreed between GO-Science and the Department, though at the end of the day it is the Department's response.

(i) Unless the follow-up action is a detailed scrutiny review, GO-Science and the Department should agree how implementation of agreed recommendations can most constructively be monitored. The discrete process of the Department formally reporting implementation progress to the GCSA is inconsistent with the partnership approach underpinning the new review model. A more constructive relationship will one based on occasional conversations between GO-Science and the Department about progress, including offers of support by GO-Science – eg pointing the Department to another Department where the issue in question has been resolved. The annual Departmental self-assessments provided to GO-Science could include a note on implementation of agreed recommendations.

(j) Where the GCSA and the Departmental Permanent Secretary agree that a detailed scrutiny review is needed, there is little value in a Departmental response other than placing on the public record the fact that a further review is being carried out.

Detailed scrutiny reviews

7. Most of the same features as in short reviews will be present. The main variations will be:

(a) The key issues for the review will already have been identified in the short review, though that should not prevent either the Department or GO-Science reaching a mutual agreement to include additional issues.

(b) Bringing the evidence base developed by other professional disciplines within the scope of the review should be considered as a matter of course.

(c) There may be a need for a larger number of external experts, depending on the planned scope of the review.

(d) The review plan is likely to be less dependent on standard metrics, though these would still have a role if the review was extending the breadth as well as the depth of scrutiny. Judgements will need to be made, in conjunction with the external experts, about the extent to which in-depth case studies should figure in the review and the

basis on which they should be selected. A well-managed and targeted stakeholder consultation may form part of the review. The maximum duration of the review should be 6 months.

(e) The assessment stage will be more drawn out than in the short review, though the external experts will only be on site at intervals. They may, however, find it valuable to spend a couple of days – including an overnight – away from the Department reflecting among themselves on the findings. There may be a need for the evidence accumulated during this stage to be documented more thoroughly than in the short reviews, but again this should be proportionate to need.

ANNEX B

EXTERNAL EXPERTS: CORE SKILLS AND ATTRIBUTES

Common to all

High status and seniority in the scientific community, as well as experience of having managed a substantial scientific enterprise such as an academic department or a major laboratory.

Recommendation by a relevant professional body (this may be solicited by GO-Science)

Ability to engage credibly and pragmatically with Departmental officials, both on specialist subjects and on relevant business processes.

For reviewing how Departments access science

Substantial expertise in a specialism relevant to a significant aspect of the Department's business

For reviewing how Departments use science

Some familiarity – not necessarily first hand experience - with the policy-making process in government.

Standing pool

Willingness to develop expertise in the new review model, and to support non-pool members in understanding it.

Available to undertake up to 4 short reviews a year, or 1 short plus 1 detailed scrutiny.

ANNEX C

SCIENCE REVIEWS: OPERATIONAL INFORMATION

Case studies

Analysis of four Reviews[28] shows that:

(a) Of a total of 91 recommendations in the main reports, 36 did not link to any case study findings.

(b) Of a total of 78 individual findings from the case studies, 37 could be held to support directly a recommendation in the final report and there were tentative, less clear links from a further 16 to a recommendation.

(c) 60% of the recommendations with direct links to case study findings related to 4 and 5 of the ten criteria. Review recommendations relating to criteria 1,2, and 3 were not supported by case study findings.

(d) Of the recommendations supported by the 37 findings, 11 recommendations were also supported by the findings of the peer reviews (see below).

Peer reviews

Analysis of the four Reviews shows that:

(e) Of a total of 91 recommendations in the main reports, 45 did not link to any peer review findings.

(f) Of a total of 35 summary or cross-project findings from the peer reviews, 9 could be held to support directly a recommendation in the final report and there were tentative, less clear links from a further 7 to a recommendation.

(g) 70% of the recommendations with direct links to peer review findings related to 4 and 5 of the ten criteria. Review recommendations relating to criteria 1,2, 6, 9 and 10 were not supported by peer review findings.

(h) Of the recommendations supported by the 9 findings, 6 recommendations were also supported by the findings of the case studies (see above).

Interviews

A total of 300 staff and stakeholders were interviewed for the four Reviews (364 if the DCMS Review is included). The total for each Review ranged from 61 (CLG) to 93 (HO/MoJ).

[28] The published reports excluding that for DCMS (which did not make use of the case study adopted in subsequent Reviews).

ANNEX D

THE TEN CRITERIA AND THE EIGHT ATTRIBUTES

This annex is background information.

The Ten Criteria

These make up the framework against which Departments are assessed a the Science Review. The following description of the criteria is provided by GO-Science:

The Science Reviews focus on the ten success criteria identified by the Chief Scientific Advisers' Committee as underpinning good practice in the use of science by Government departments. These relate to how effectively departments:

1. **Develop a clear, overall science strategy**

Departments should take a strategic approach to setting R&D budgets, and should publish science (or evidence) and innovation strategies that set out the broad framework within which research programmes and other science-related activities are carried out. This is an important step in linking research and development to the effective delivery of a department's objectives and show how value for money is achieved.

2. **Horizon scan – to identify future science-related issues**

Horizon scanning is defined as the systematic examination of potential threats, opportunities and likely future developments, which are at the margins of current thinking and planning. Horizon scanning may explore novel and unexpected issues, as well as persistent problems or trends. Departments should regularly undertake horizon scanning to improve the robustness of their evidence base and policies.

3. **Review and harness existing science and identify gaps and opportunities for future research**

To demonstrate value for money and effective use of resources, departments should have in place effective arrangements for deciding what current or potential science could benefit the department's delivery of its objectives and hence whether new research is needed or where it would best be targeted. In particular, departments should actively manage existing knowledge, synthesise existing research, and work with Other Government Departments (OGDs) and the research bases in the UK and internationally.

4. **Commission and manage new science** [and]
5. **Ensure the quality and relevance of the science they carry out and sponsor**

As part of the drive for evidence-based policy and improved service delivery the Government needs to use, and be seen to use, high quality science and the most appropriate new technologies. Science programmes funded by Government departments make a very important contribution to policy formulation. Even though the outcomes of the science itself cannot always be predicted, departments must be able to commission the right science, assess its quality, and use it effectively. The credibility of departmental policy-making generally will be undermined if individual policies are perceived to be based on poor, or the wrong science.

6. **Use science and scientific advice**

Departments need scientific advice to underpin their policy making and regulatory

activities. Such advice can be provided by external or internal experts, and / or informed by the output of research programmes commissioned by the department. There needs to be an effective communications bridge between the experts and the policy makers.

7. **Publish results and debate their findings and implications openly**

In accordance with the Freedom of Information Act and to ensure robust interpretation of scientific findings and their policy implications, departments should publish and openly debate scientific results.

8. **Share, transfer and manage knowledge**[29]

Knowledge transfer should be treated by departments as a strategic goal and enjoy high-level focus.

9. **Follow the Guidelines on Scientific advice and policy making**[30] **and the Code of Practice for Scientific Advisory Committees**[31]

The Guidelines on Scientific Advice and Policy Making provides high-level high level guidance and were used in formulating these ten key criteria for the Science Reviews. Its key messages are that departments should:

i. think ahead and identify early the issues on which they need scientific advice;

ii. get a wide range of advice from the best sources, particularly where there is scientific uncertainty; and,

iii. publish the scientific advice and all relevant papers.

The purpose of the Code of Practice is to provide more detailed guidance specifically focused on the operation of scientific advisory committees and their relationship with Government and to help them translate the principles in the Guidelines into day-to-day practice.

10. **Use, maintain and develop scientific expertise (including both capacity and capability building)**

Whether a department has its own dedicated research unit, or commissions work from outside organisations, it needs to ensure it has long-term access to experienced scientists who are able to understand and interpret issues at the science-policy interface, taking into account the full range of scientific opinion as appropriate.

8. The ten criteria are used to identify both areas of good practice, especially those that could be adopted elsewhere, as well as areas for improvement.

[29] Including exploitation of intellectual property where appropriate.
[30] Office of Science and Technology. Guidelines 2005: Scientific advice and policy making, October 2005
[31] Office of Science and Technology. Code of Practice for Scientific Advisory Committees, December 2001

The Eight Attributes

Alongside the 2004 Spending Review, the Government published the Science and Innovation Investment Framework 2004-2014[32], commonly referred to as "the ten-year framework". The framework sets out eight attributes for the effective management of science and research across government, to be achieved by 2014. These are:

1. The Government as a whole, and all Government departments, will have adopted a culture of using sound scientific advice to inform policy development, delivery and departmental decision-making. This should involve DCSAs in all major departments with direct access to Ministers and departmental managers, and with departmental managers involving DCSAs on all major policy issues, not just those with obvious scientific aspects.

2. All scientific work commissioned and used by Government will be of appropriately high quality, drawn from the best possible sources (including the science base and the private sector), commanding the confidence of Government Ministers and officials. Government departments will be paying the full economic costs of the research they commission from universities.

3. Priorities for research will be set at the strategic level, not just within departments as they are now, but also across government as a whole, taking account, for example, of the 2003 Strategic Audit from the Cabinet Office. CSAs – acting as a group – along with other bodies, such as the Council for Science and Technology, will provide advice on the prioritisation of strategic issues. The use of science in policymaking will be applied consistently across the board where an issue affects more than one department.

4. All Government departments will be using sophisticated scientific horizon-scanning techniques, linked both to their own policy horizon scanning, that of other departments, and to the OST horizon-scanning centre. This should involve departments drawing upon the science base to ensure they are informed about future risks and opportunities. Cross-departmental science initiatives, such as the Foresight programme and Prime Minister's Strategy Unit work, should develop and disseminate best practice guidelines and should provide capacity to deal with selected issues, working closely with other departments.

5. Scientific expertise will be used to the maximum effect possible, allowing greater use of Research Council, charity and private sector input to Government advice, and giving Government scientists greater opportunities to contribute to the work of the science base and the exploitation of their work in the wider community, industry and commerce. Analysts, including scientists, will be able to network more effectively – within their own department, across departments, Research Councils, the private sector and internationally – to ensure awareness not just of research results already generated but also active research underway elsewhere.

6. Knowledge transfer objectives will be fully incorporated into departments' S&I strategies, and scientific advice on procurement in Government departments will be seen as a natural and logical means of pulling through the development of new technologies.

7. The use of scientific knowledge will have been fully integrated into

[32] HM Treasury, July 2004

Government analytical and risk assessment processes, and risk assessment guidance will be consistent with the advice in Guidelines 2000. Science will be regarded as one of the key analytical inputs to decisions along with specialisms like economics, law and statistics, with policy staff at all levels aware of the need to seek scientific advice – in the same way as they incorporate economic and legal advice.

8. Scientific advice for the Government will be generated in a fully inclusive manner and command the support of the public and other stakeholders. Scientists, including Government scientists, will have the training and willingness to communicate openly with the public, including through the media. Politicians and the public will understand what science and research can and cannot deliver, in particular that science and analysis will provide information and knowledge to those who must take decisions, but that it is for politicians and for the public to take the decisions themselves.

Although delivery of the eight attributes is to be achieved through a variety of strategies and plans, GO-Science monitors progress through an annual self-assessment questionnaire.

ANNEX E

SUMMARY OF ASSURANCE PROCESSES OPERATED BY OTHER PROFESSIONS

Government Economic Service (GES)

The GES operates a programme of a peer reviews of Departmental economics activities. The reviews are intended to help Departments address two key issues: making better use of economics, and making economists better. They provide an analysis based on assessments by fellow Government Chief Economists and other stakeholders, both within the Department and beyond; and suggest actions which the Departmental Chief Economist(s) can then take forward in consultation with their Permanent Secretary and the Heads of the GES.

The reviews are planned to last for no longer than 12 weeks, from starting to establish the review team to presentation of the agreed final report. The fieldwork is concentrated in a 4–week period in the middle of the review. GES expect reviews to be shorter for small Departments, and longer if the number of external stakeholders is very high. The GES Board of Chief Economists is aiming to achieve a lighter procedure for future reviews.

Because the performance of individual economists or teams can form part of the review, reports are not published.

Government Operational Research Service (GORS)

GORS does not currently operate a centrally run review process. Departments carry out their own assessments of OR work as necessary, and informal peer reviews can form part of this. Staff development and professional standards are secured using the GORS Competency Framework which includes the Professional Skills for Government framework and reference material for a large number of OR specific techniques.

Government Social Research service (GSR)

The GSR are partners with GO-Science in the Science Review programme, although the reviews are managed and staffed by GO-Science. A senior member of the GSR normally sits on Science Review Steering Panels.

GSR has developed an addendum to the Civil Service code which sets out the principles underpinning the quality of GSR Members and the work they produce. The 'GSR Code' also acts as a framework for assessing social research capability in member departments. The profession is introducing a quality assurance process which will be based on a combination of self-assessment by the Head of Profession and an independent peer review. The self-assessment process has already been piloted and used to identify where practice can be shared across departments, and where new guidance is needed. The peer review process is in development, with the aim of starting the first reviews in 2008/09. These will be

short (certainly no more than 12 weeks), covering up to six departments per year; and each review panel will involve social science experts outside the civil service, as well as GSR Heads of Profession from other departments. The intention is to publish the key findings from the independent peer reviews.

Government Statistical Service (GSS)

Co-ordinated by the Office of National Statistics the GSS has operated a specific quality review programme for National Statistics since 2000. This has led to marked improvements in the fitness for purpose of those statistics subjected to review. Most reviews were completed in six to nine months, though others took between 12 and 24 months. Three categories of review – strategic, standard and targeted – were operated from 2002.

A review of the programme recommended several changes in the light of experience, including a closer alignment between the review schedule and the Spending Review cycle, and the introduction of two-tier system of reviews (strategic, and Departmental in-house).

Under the terms of the Statistics and Registration Service Act (2007), the new UK Statistics Authority is required to assess official statistics against its Code of Practice, and either designate them, or remove their designation, as National Statistics. The Statistics Authority is currently developing its thinking about assessment, and the ONS will ensure that a refocused Quality Review Programme is consistent with the provisions of the new Code of Practice, and hence the assessment function.

ANNEX F

REVIEW TERMS OF REFERENCE, GOVERNANCE AND METHODOLOGY

Terms of reference

The review's terms of reference were as follows:

To carry out a high-level review of the Government Office for Science's Science Review Programme and make recommendations for taking the Science Review Programme forward, particularly around how quality assurance across government's analytical disciplines could be joined-up.

The Review will look at and assess:

- how effective the Programme has been to date in meeting its overall aims, which are:
  - maintain and improve the quality and use of science in government;
  - review existing departmental systems for assuring the quality, management and use of their science; and,
  - disseminate examples of good practice from within the UK and abroad
- how cost-effective the Programme has been has been;
- the strengths and weaknesses of the Programme's approach and methodology.

The Review will compare how the Science Review Programme relates and compares with Government's assurance mechanisms for the other analytical professions (i.e. economics, statistics and operational research) and other relevant reviews of how government departments work (e.g. the capability and OGC gateway reviews).

The Review will make recommendations about how the Science Review process could be improved; simplified; streamlined; joined-up with the quality assurance processes for the other analytical professions; and linked to other Government Reviews (e.g. the capability reviews). Recommendations will also consider the associated potential benefits and risks (e.g. balancing pace with collection of sufficiently robust evidence base). Recommendations may address any aspect, but will include:

- the scope for thematic as well as (or instead of) departmental reviews;
- the scope for shortening the public consultation part of the Review methodology;
- how to balance the use of independent panel members, peer reviews, consultation and case studies with desire for a faster pace;
- how to balance the benefits of making a small number of targeted recommendations against the need to be comprehensive, covering all aspects of how a department does science/analysis;
- the scope for moving towards inclusive, "Evidence and Analysis Reviews" covering more (or all) of the analytical professions; and
- how the Reviews could better link to the Capability Reviews in terms of timing and remit.

The Review will Report in May 2008

Governance

The review has been guided by a steering group made up of the heads of the government analytical professions (the Heads of Analysis group) augmented by

two Departmental Chief Scientific Advisers and officials from GO-Science (marked with an asterisk). The full list is:

‡Professor John Beddington, Government Chief Scientific Adviser, and Head of the Government Office for Science

*Judy Britton, Joint Head of Science in Government, Government Office for Science

*Jeremy Clayton, Deputy Head of the Government Office for Science

*Professor Brian Collins, Chief Scientific Adviser, Department for Transport and Department for Business, Enterprise and Regulatory Reform

*Professor Sally Davies, Chief Scientific Adviser, Department of Health

Karen Dunnell, National Statistician, and Chief Executive, UK Statistics Authority

‡Nick Macpherson, Permanent Secretary, HM Treasury

Tony O'Connor, Chief Government Operational Researcher, HM Treasury

Vicky Pryce, Joint Head of the Government Economic Service, and Director-General, Economics, Department for Business, Enterprise and Regulatory Reform

‡Dave Ramsden, Joint Head of the Government Economic Service, and Managing Director for Macroeconomic and Fiscal Policy, HM Treasury

Professor Paul Wiles. Government Chief Social Researcher, and Chief Scientific Adviser, Home Office

Siobhan Campbell, Government Social Research Unit, HM Treasury (Secretary to the Heads of Analysis Group)

*Patrick Erwin, Head of Science Review Team, Government Office for Science (Secretary to the Review steering group)

The steering group has been chaired variously by those marked (‡). Meetings were held on 18 February (to agree the terms of reference), 17 April (to review emerging findings and provisional conclusions) and 6 May (to review a draft report with recommendations). The final draft was seen by the Chief Scientific Advisers' Committee on 5 June, and reviewed in correspondence by the steering group at the end of July.

Methodology

The majority of the fieldwork was carried out during March and April 2008:

- 40 discussions, face-to-face or by telephone, with a selection of:
  - Departmental officials, both with and without direct experience of Science Reviews
  - Members of Science Review Steering Panels
  - Stakeholders who were interviewed as part of a Science Review
  - Officials of the Government Office for Science, including the Science Review Team

- Examination of documents, including a selection of those relating to individual Science Reviews and to the management of the Review programme as a whole.

In reaching conclusions, I have focussed principally on the 5 reviews completed to date. The first review – of DCMS - adopted a methodology which differed in some respects from the subsequent reviews. For this reason, and because the first attempt at a new process inevitably has the characteristics of a pilot, the DCMS review has been given less weight than the others in this assessment. I have also held discussions with the Review Team, Departmental officials and Steering Panel members for the Department of Health review which is in its final stages; and for the Food Standards Agency, which has only recently started, with the Review Team and and the Agency's Chief Scientist. .

In identifying sources to inform the review, I have aimed for representativeness rather than comprehensiveness. This is partly because of time limitations. More significantly it is because a review of this sort is as much about perceptions as about hard facts. Credibility is central to an external scrutiny process of the sort the Science Review programme seeks to be, and its presence or absence has a major impact on the effectiveness of the scrutiny.