Office for Environmental Protection

Paper 24.13

Board Paper

Date

12 February 2024

Title

Measuring our success – a review of our performance framework

Report Author

Andy Lester, Head of Business Strategy and Planning

Responsible Executive Director

Richard Greenhous, Chief of Staff

Paper for decision

Open paper

Issue

1. Our performance framework has not developed as much as intended since the Board agreed it in May 2022. We committed to have further developed it by our strategic review. This paper proposes how to evolve our approach.

Recommendation

2. The Board is recommended to agree;

    a. that we should develop a balanced scorecard approach to reporting our performance, with the core components outlined in this paper (subject to further development of detail where outlined).

    b. that it should consider the balanced scorecard quarterly, and a qualitative report on performance and impact half yearly

    c. that we should publish:

        i. the balanced scorecard in our annual report and accounts, alongside case studies and wider narrative as appropriate

        ii. an interim performance update, including the balanced scorecard, alongside our annual report and accounts

iii. a version of each evaluation study we commission

Background

3. The NAO has an extensive body of work on performance measurement. It argues that it is essential for public bodies to measure and report transparently on their performance so that they are more likely to achieve their objectives, and can be held accountable for doing so by Government, Parliament, and the public.

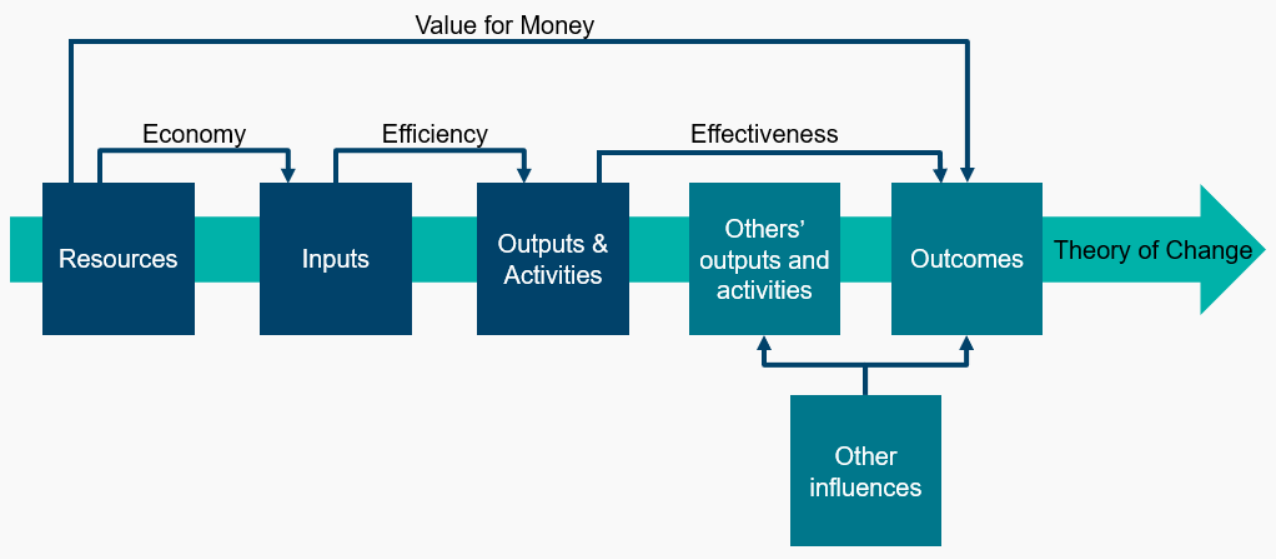
4. Performance frameworks build on concepts such as theories of change or logic models. These aim to identify the causal links between the effort an organisation applies, its activities, and the outcomes or impact they achieve. In a perfect system, an organisation would assess:

   a. Whether intended interim and long-term outcomes are being achieved (impact)

   b. Whether outputs or activities are causing those outcomes to be achieved (effectiveness)

   c. The total resources needed to generate those outputs (efficiency)

   d. How well the application of resources (efficiency) is leading to outcomes (effectiveness) – value for money.



5. All organisations face tension between measuring what can be measured (typically resources, inputs and outputs) and what is most insightful (typically impact, effectiveness and value for money). The NAO argues that measuring effectiveness is particularly complex for oversight bodies as the relationship between outputs and outcomes is mediated by the actions of those we oversee, and time horizons are often long.

<u>Our current performance framework</u>

6. In our strategy we committed to develop a structured approach to assessing our impact, building over time from an initial framework the Board agreed in July 2022. Each of our corporate plans has similarly set out an aim to further develop how we measure our success.

7. Our current performance framework largely only includes measures of the volume of activity we undertake – complaints, enquiries, advice, reports. It does not seek to link these to resources (to show our efficiency) or outcomes (to show our effectiveness). It is at the lower end of the NAO's maturity model for performance management. It exists, and broadly functions. It does not enable, challenge or optimise our work.

8. For any organisation, this narrow focus limits the utility of the information to ourselves and others. The nature of our strategic, targeted work is such that it will always be of a low volume, making metrics of outputs of even less value as greater volume is not a marker of success. Instead, we rely on the few activities we choose to do to be influential.

9. We capture evidence from specific cases of the influence we have had. We include these in our annual report, and publish examples of our casework on our website. We periodically consider the extent to which we have achieved our intent in specific projects. We have also undertaken limited external research on our effectiveness, notably through our research into stakeholder perceptions. We also monitor coverage of the OEP in the media and Parliament.

10. In discussions, the Board has expressed concerns on evolutions to this framework, including:

    a. scepticism in our ability to simplify the complex picture of our influence, into measurable activities or indicators

    b. concerns over perverse incentives of some of the more likely measures we could choose, and

    c. some concern of the self-selecting nature and lack of objectivity of a case study led approach.

<u>Reporting our performance</u>

11. We report performance metrics to Defra and DAERA quarterly. We know that improved performance metrics have been a focus of interest for previous Ministers at times – such interest is likely to return.

12. Our annual report includes narrative and qualitative description of what we have done. In auditing our two reports to date, the NAO have questioned the absence of performance data to underpin and provide an objective balance to our narrative, and the absence of description of the impact we judge we have had, rather than what we have done. More is expected to meet good practice expectations, and to live our principle to be as transparent as we reasonably can.

13. We provide aspects of our performance measures – notably our delivery to plan – routinely to support decision making, both as an Executive and Board. Our wider performance measures, in current form, have not proved particularly helpful in managing the organisation.

14. Yet, it is our strategy to make the most difference and to prioritise those things which we judge to provide the greatest capacity for us to do so. We require every project to set out its strategic intent. We could have more structure to our assessment of the extent to which we achieve what we intend.

Analysis

15. One model recommended in guidance is to adopt a 'balanced scorecard' of performance, so as to present different dimensions or perspectives of performance concurrently. This is said to support organisations to have broader thinking and make the trade-offs between dimensions of performance more transparent.

16. Balanced scorecard models typically include metrics on four dimensions: delivery, financial, organisational development/learning and customer. In our context our 'customer' is those we seek to act or change behaviour in light of our work. At different times this might be Parliament, the Assembly, Government and Public Authorities and possibly the public. Our influence – the 'right hand side' of the logic model set out above – is therefore our most appropriate approximation of customer.

17. We recommend a balanced scorecard for the OEP which seeks answers to some or all of these questions, considered in detail further below.

Delivery • What outputs have we delivered? • Will we deliver what we have promised?	**Influence** (customer) • Have we been heard? • Have our scrutiny outputs had influence in Parliament and the Assembly? • Have our scrutiny outputs had influence in Government? • Have our compliance activities helped resolve issues identified?
Finance and Resources • Do we efficiently create our outputs? • Do we have the people we need? • Do we meet our service standards? • Are we compliant and well run? • Are we sustainable in our use of natural resources?	**Organisational excellence[1]?** (growth) • Have we engaged and listened to others? • Have we engaged and motivated our people? • Have we grown our expertise? • Are we developing organisationally as we intend?
Overall 18. How do others judge us?	

<u>Measuring our delivery</u>

19. Our existing framework captures information on the outputs we deliver e.g., the number of complaints and enquiries we handle, the investigations we launch, the advisory activities we complete, and the reports we publish.

[1] In our current strategy, this would approximate to do we have the relationships, authority, expertise and voice we need to succeed?

20. We also measure the extent to which we will deliver our corporate plan. Little further development seems required.

21. Potential elements to the balanced scorecard are set out below. Aspects to draw to the Board's attention are in red text.

Indicator	**Type**	**Reported as / to**
# of reports laid before Parliament and the Assembly*	Existing Output Metric Objective 1, 2, 3	Metric Public / Gov't / Board Quarterly
# of advisory activities (advice, consultation responses, evidence to Parliament)*	Existing Output Metric Objective 1, 2, 3	Metric Public / Gov't / Board Quarterly
# of complaints received*	Existing Input Metric Objective 1, 2, 3	Metric Public / Gov't / Board Quarterly
# of enquiries received*	Existing Input Metric Objective 1, 2, 3, 4	Metric Public / Gov't / Board Quarterly
# of investigations launched and completed*	Existing Output Metric Objective 3	Metric Public / Gov't / Board Quarterly
# of information and decision notices*	Existing Output Metric Objective 3	Metric Public / Gov't / Board Quarterly
# of environmental reviews, urgent judicial reviews of interventions in judicial review*	New Output Metric Objective 2, 3	Metric Public / Gov't / Board Annual
% of corporate plan commitments on track or complete	Existing Output Metric Objective 1, 2, 3, 4	Metric Public / Gov't / Board Quarterly

* metric reported separately for England and Northern Ireland

Measuring our influence

22. We have to date gathered case studies as we judge we have been influential. We monitor the extent to which we are referred to in Parliament, and in the media, though not in a structured or consistent way to allow comparison. We do both these things focused on specific activities – whether we have caused a response, and the response we intend. At an organisational level, as a whole, we do not bring this together – other than through periodic reflective discussion, such as the Board had in October.

23. In our scrutiny and advisory functions, we have influence through what we say, report and recommend – directly, by our recommendations being adopted, or indirectly by shining a light on issues, and others amplifying and building on them to encourage them to be adopted. Simplified, therefore, for our scrutiny functions to be effective we need a) to be heard, b) for our recommendations to be acted upon directly, or championed by others and c) correct – so that our recommendations do lead to improvement.

24. In our compliance functions, we aim to remedy a serious failure in law we identify. That can be by changing practice or behaviour, or by clarifying and therefore improving implementation of the law. Our interest can change behaviour in and of itself. Our casework, investigations and enforcement ultimately lead us to make findings and recommendations we expect to be followed.

25. We propose to:

    a. Systematically measure the extent to which our outputs are cited in Parliament, and receive media prominence – making what we do now comparable over time, and between activities. It could be possible, but do not yet propose, to extend this to social media and academic publications.

    b. Possibly, measure the 'sentiment' attached to these citations – negative, neutral, supportive –to capture whether others judge them to be 'right'. Whilst subjective – both in classification, and because others' sentiment will be informed by their agenda and views – it is an early indicator. Further work is needed to ensure we can sustainably operationalise this.

    c. Measure the extent to which recommendations from our scrutiny and advice are adopted, recognising that this will be judgmental and may vary over time. We propose to develop a RAG model for whether a recommendation is adopted, partially adopted, not adopted or not yet known. We have done such an assessment on Government's response to our Environmental Assessments report, and can use this as a basis to develop guidance to support consistency in judgment. This is likely to require us to measure whether our recommendations are accepted, and adopted. Whether public authorities agree to do as we recommend, and whether they then actually do.

    d. Similarly measure whether our recommendations from 'compliance activity' are adopted. We could apply this equally to casework as investigations.

    e. Consistently assess the extent to which a project has achieved its intent.

    f. We could, but do not propose to, measure the extent to which complainants are satisfied with our work. We are not an ombudsman and our purpose is not to find resolution for the complaints brought to us – as our website makes clear. The looser connection between a complaint and those matters we pursue make this a less relevant and potentially misleading indicator.

    g. Through case studies continue to assess the extent to which specific outputs have had influence.

    h. After sufficient time, commission others to evaluate whether our work in a given domain was influential. This could include whether recommendations were adopted and (importantly) right for a given project, or could look at a number of outputs and activities across a programme and assess whether they supported change in that domain. This is likely to be through a commissioned study, on a retrospective basis, after a sufficient period of time has elapsed.

26. Potential elements to the performance framework

Indicator	Type	Reported as / to
Number of citations in Parliament and the Assembly (including select committees) parliamentary reports or meetings	New Effectiveness Indicator Objective 1, 2, 3	Metric Public / Gov't / Board Quarterly
Number of mentions of our outputs in monitored media	New Effectiveness Indicator Objective 1, 2, 3	Metric Public / Gov't / Board Quarterly
Sentiment of citations and media mentions (to be developed)	New Effectiveness Indicator Objective 1, 2, 3	Metric Public / Gov't / Board Quarterly
Adoption of recommendations RAG (to be developed)	New (but committed to) Effectiveness Indicator Objective 1, 2, 3	Qualitative metric/report Public / Gov't / Board Twice a year
Achievement of our strategic intent RAG (to be developed)	New Effectiveness Indicator Objective 1, 2, 3, 4	Qualitative metric/report Board Twice a year
Case studies of influence	Structure to existing Effectiveness Objective 1, 2, 3, 4	Qualitative report Board Twice a year
Evaluation studies	New (but committed to) Effectiveness Objective 1, 2, 3, 4	Qualitative report Board Every other year from 2025/26

Finance and Resources

27. We monitor our expenditure, allocation of people and non-pay expenditure closely. It is instrumental to managing our business and delivery. Existing data could be adapted to readily provide a view of performance which could be useful to stakeholders, such as:

    a. The extent to which we spend the resources Parliament voted to us

    b. The extent to which we have people in post to deliver

    c. The extent to which we are mitigating our environmental impact

    d. The extent to which we meet our service standards in our public facing services (FOIA, enquiry response times, etc) – we can measure the timeliness of our responses through existing data, if we were to choose to, and draw comparisons over time. This could be extended to include complaints, though there is likely too much volatility in these response times, and too much influence of external actors, to make this meaningful – even in preliminary stages, such as validation checks.

    e. The change over time in the cost 'per unit' of some of our activity – though this would require us to enhance the data we collect about actual costs incurred, including through time spent. We could do this on a systematic basis, or ad-hoc – where we wish to understand whether the resources we chose to apply yielded the impact we intended.

28. Potential elements of the performance framework

Indicator	Type	Reported as / to
% expenditure forecast to allocated resources	New Efficiency Indicator Objective 4	Metric Public / Gov't / Board Quarterly
% staff in post (total planned OD less vacancy)	New Input metric Objective 4	Metric Public / Gov't / Board Quarterly
% sickness absence	New Input metric Objective 4	Metric Public / Gov't / Board Quarterly
% change per FTE in carbon emissions, water consumption and waste generation vs baseline	New Input metric Objective 4	Metric Public / Gov't / Board Annual from 2025
% FOIA and EIR responses within statutory timelines	New Efficiency Indicator Objective 4	Metric Public / Gov't / Board Quarterly
% enquiries responded to within XX days (to be developed)	New Efficiency Indicator Objective 4	Metric Public / Gov't / Board Quarterly
Unit cost studies	New Efficiency Objective 1, 2, 3, 4	Qualitative report Board Every 2nd year from 2025

Organisational Excellence

29. Our existing strategy identifies certain characteristics of excellence as critical – relationships, authority, expertise and voice. Our draft guiding policy continues to see these as essential, along with trust and purpose. Our strategy and plans also set out ways in which we aim to develop our capabilities. For example, by developing our approach to analysing the trajectory of improvement.

30. We could measure:

    a. The extent to which we engage with stakeholders. We capture data now, it could be organised and reported differently as a measure of performance – this could focus on the volume of engagement, how this is distributed in different stakeholder groups to show breadth (NGOs, public authorities, business groups, local authorities, etc), or new connections made. It could also be expanded to include stakeholder satisfaction from events, where we seek such feedback.

    b. The engagement of our staff – as we do now. There might be other useful measures to draw on from our staff survey, such as the proportion of staff who have what they need to do their job well.

    c. The extent to which we develop specific capabilities within our strategy – this is likely to be 'delivery' measures around specific projects, such as in our current plan those related to baseline environmental law monitoring, and analysing the trajectory of environmental improvement – this would be a subset of our delivery tracking, focussed on whether we have delivered particular parts of our plan

d. The extent to which we access a breadth of expertise we need, and/or the extent to which we make such evidence publicly available. Breadth is difficult to measure – a proxy could be how many instructions through our College of Experts we make, how many independent research reports we commission and publish.

31. Potential elements of the performance framework

Indicator	**Type**	**Reported as / to**
People survey engagement metric	Existing Efficiency Indicator Objective 4	Metric Public / Gov't / Board Annually
People survey equipped for role or equivalent (to be developed)	New Efficiency Indicator Objective 4	Metric Public / Gov't / Board Annually
People survey participation metric	Existing Efficiency Indicator Objective 4	Metric Public / Gov't / Board Annually
# of engagement activities (to be defined)	New Efficiency Indicator Objective 4	Metric Public / Gov't / Board Quarterly
% of 'organisational development' corporate plan commitments on track or complete	New Output Metric Objective 4	Metric Public / Gov't / Board Quarterly
# of engagements with College of Experts	New Output Metric Objective 1, 2, 3	Metric Public / Gov't / Board Quarterly
# of evidence reports published	New Output Metric Objective 1, 2, 3	Metric Public / Gov't / Board Quarterly

<u>Overall</u>

32. Alongside our existing measures of delivery, we set out in our strategy and corporate plans that we would seek regular and active feedback on our performance, including by understanding how others judge our effectiveness and influence. To do so, we envisaged periodic evaluations such as:

a. In 2022, we undertook, but did not make public, research on stakeholder perceptions of the OEP – our positioning and voice. That provided a qualitative baseline of how we were judged in our early years. It is wise to repeat such research periodically – every 2 to 3 years.

b. We also judged that in time, we could commission independent research into our effectiveness – overall, or in specific programmes or priorities such as in 24.h above.

c. We have baseline evidence (from January 2022) of public awareness of the OEP, through the research YouGov undertook on our behalf. We could undertake further research. Public awareness is not an end we have yet pursued – and it seems unlikely that we would learn from the findings. If our strategy evolves to focus more on public awareness, this might be more warranted.

33. Potential elements of the performance framework

Indicator	Type	Reported as / to
Stakeholder perceptions survey	Existing (but to repeat) Effectiveness Objective 1, 2, 3, 4	Qualitative report Board Every other year from 2024

Transparency / Reporting

34. We publish a quarterly complaints report, whose purpose is to be informative about our complaints rather than make transparent our performance in managing them. We publish information in our annual report and accounts and case studies of our casework.

35. It is our strategy to be as transparent as we reasonably can, because we have judged our credibility and authority to act to depend on public confidence in the judgments and decisions we make. There are reasons of accountability and good governance to be transparent, particularly for independent organisations over which Parliament, Ministers and therefore the public have fewer levers of control.

36. We propose to:

    a. Publish a scorecard with consistent measures (as we agree) within our annual report and accounts.

    b. Publish a mid-year scorecard of performance, outside of, but to align with, the annual report. This is the approach some of Defra Group's delivery bodies take for example. There are advantages to doing so – for example, it would allow us to explain progress against our corporate plan, including by articulating how and why we have made different in-year prioritisation choices in light of events that arise.

    c. Publish the evidence we create of the views others have of us. We did not publish the report we received on stakeholder perceptions, but did publish YouGov's research. It is likely that we could in future commission all research in a way that would be internally and publicly useful.

37. There are different ways of presenting information. For example, it may be that we would not choose to publish a metric about the extent our recommendations are adopted, but instead an annual qualitative evaluation of whether and how they have been so as to situate a metric in broader context – given the subjective content.

38. We could, but do not propose to publish regular information on aspects of our performance. NAO for example have a live 'recommendations tracker' where they maintain a log of recommendations, and the extent to which they are adopted. It is transparent, but granular and not immediately useful in judging NAO's performance.

Using the framework

39. Literature and guidance on performance management includes extensive examples of the risks of poorly identified indicators, the over-reliance on individual or small groups of indicators and the perverse incentives from single targets for example. Each of these speak to the inappropriate influence of measures on management decisions, and action.

40. The Board has had reservations as to the perverse incentives that could arise from measuring the extent to which our recommendations are adopted. There is concern that we will recommend what will be agreed, rather than what is needed, and achieve less for environmental protection and improvement as a result.

41. Perverse incentives exist with any measure that distils the complex into the manageable, and oversimplifies reality as a result. In capturing any measure, it is arguably unavoidable that there are perverse incentives. Key is how those perverse incentives are managed.

42. Perverse incentives are only damaging if they are allowed to influence decisions and behaviour. For a measure to drive poor behaviour there must be a failure of leadership, decision-making and culture – so that the measure overrides other, wiser, more rounded judgments about what is right and what is needed. Given the collective nature of most decisions, and limits to individual's delegated authority to act, in our structure, those would be failures of the Executive Committee and the Board's judgments. It is hard to envisage that failure of governance, in our context and current culture.

43. In this example, it is also undeniable that the extent to which our recommendations are adopted tells us *something* about the effect of our work, even if it does not tell us *everything* about the effect of our work. It seems we could have the organisational confidence that we will use information we gather wisely.

44. We propose to implement and report on the framework as follows:

Products	**Audience**	**Frequency & Vehicle**
Performance dashboard	Executive Board Defra/DAERA Public	Quarterly Quarterly CEO report annex Quarterly report Annually (AR&A) Mid-year performance update (interim update)
Performance report (including recommendations adopted, compliance resolved, case studies, and review of strategic intent)	ExCo / Board Public	Twice yearly Excerpts annually (within AR&A)
Stakeholder perception report	Board Public	Every other year (from 2024)
Evaluation study	Board Public	Every other year (from 2025)
Unit cost study	ExCo Board	Every 2nd year (from 2025)

Northern Ireland

45. The framework would apply equally to England and Northern Ireland.

Finance and Resource

46. There are incremental resource demands associated with these proposals.

47. Almost all the data proposed to be collated within the performance report exists. Further analysis of the time cost of transposing that into a consistent format, and quality assuring it for any wider use is needed, subject to the Board's steers on content sought in this paper.

48. Once established, it is not envisaged that this would be material – perhaps in total 1 day a quarter, after the original work to ensure it is well designed.

49. Resources are required to develop an approach to assessing the extent to which our recommendations are adopted, whether we achieve our strategic intent, and to make precise and establish the data to support the metrics proposed. Resources would then be required to produce the proposed reports:

    a. The performance and delivery report summarising these components, and gathering case studies for ExCo and the Board to consider

    b. The external mid-year performance update

    20 days time is included in the BS&P business plan to develop this approach, including for consultation, and then to operate in the first year. This is around 0.15 FTE.

50. There are costs associated with commissioned activity. In 2022, we spent £37k on the research report Woodnewton Associates completed for us. We assume £40k and around 20 days. A more detailed evaluation study is likely to be more expensive; we assume costs similar to an evidence project – c£80k, and around 40 days.

51. For 2024/25, the affordability of these proposals relative to other priorities is uncertain. These will be highlighted to the Board in its planning decisions in due course.

Impact Assessments

Risk Assessment

52. There is a risk that our existing framework is insufficiently mature to secure the confidence of the public, government and others, and that the NAO and others will judge our approach to be inconsistent with expectations of public sector reporting and good practice. These proposals seek to further develop our proposals in mitigation of this risk.

53. There is a risk that limited further development of our framework will not meet the legitimate expectations of the public and stakeholders, given the commitments made at the time of our Strategy and each of our Corporate plans.

54. There is a risk that the indicators, and active reporting of them, distorts organisational priorities and creates unintended consequences. This is mitigated through the Executive Committee and the Board's influence on the interpretation and use, including that all key recommendations are reserved to the Board.

55. There is a risk that the burden of collection and presentation is disproportionate to the value of the products. This is to be monitored ongoing with adjustments made as required. The proposals do incur resource, and must be sustained once we begun.

Equality Analysis

56. No material equalities' impacts have been identified.

Environmental Analysis

57. This framework aims to identify ways to assist the OEP to understand extent to which our activities are contributing to its strategy and principal objective.

Implementation Timescale

58. We are committed to have developed our performance framework by the time of our next strategic review. We intend to develop this proposals for inclusion in any consultation.

59. We would then propose to operate these proposals during consultation in H1 of 2024/25 so as to learn further, and so that learning can inform the approach we adopted.

Communications

60. This approach would be communicated externally within the communications strategy to support our strategy review, both its consultation and adoption.

External Stakeholders

61. We have discussed the proposals within this framework with Environmental Standards Scotland. Our proposals are in similar to the approach ESS has developed, adapted and extended to our context.

62. We would propose to engage with stakeholders within our strategy review, as is appropriate to the issues prioritised for that engagement. ESS consulted on their performance framework in 2022/23. They report that there little to no comment on the proposals presented.

63. These proposals have been informed by discussions and information provided by Malcolm Beattie in 2023.

Paper to be published	Yes
Publication date (if relevant)	With meeting minutes
If it is proposed not to publish the paper or to not publish in full please outline the reasons why with reference to the exemptions available under the Freedom of Information Act	Elements of this paper may be redacted as publication would harm the effective conduct of public affairs, including the Board's ability to receive candid advice and engage in free and frank discussion (s.36)

(FOIA) or Environmental Information Regulations (EIR). Please include references to specific paragraphs in your paper	