Final Report
October 2020

Evaluation of the Defra Complexity Evaluation Framework (CEF)

steer economic development

Final Report
October 2020

Contents

1   Introduction

1.1   This report outlines the key findings from the Evaluation of the Complexity Evaluation Framework (CEF) carried out by an independent evaluation team from Steer Economic Development (Steer-ED).

1.2   The CEF is an evaluation resource, produced in conjunction with and for the use of the Department for Environment, Food and Rural Affairs (Defra) by the Centre for the Evaluation of Complexity Across the Nexus (CECAN)[1]. The CEF was published in January 2020[2].

1.3   CECAN has worked with Defra over the past two to three years to bring complexity thinking into the department as part of its wider work in building capacity in evaluation methodologies within government departments. It's commercial arm, CECAN Ltd, was commissioned by Defra's Strategic Evaluation Team (SET) to develop the CEF.

Introduction to the CEF

1.4   As outlined in the Request for Quotation (RfQ) for this evaluation, the CEF is *'a toolkit intended to inform the commissioning of evaluation across the remit of the Department for Environment, Food and Rural Affairs'*, which will *'guide analysts and policy makers within Defra, as well as external evaluators, in ensuring complexity thinking is appropriately embedded into evaluation design and methods'*.

1.5   The CEF is divided into two parts. The first part introduces the concept of complexity, explains its relevance to Defra, and gives examples of some of the main characteristics of complex systems which are relevant to Defra policy areas. The second part is a framework of considerations for those commissioning and managing evaluations in complex contexts. It includes three sections, reflecting the major interconnected phases of an evaluation:

- *Understanding*: determining the evaluation purpose, understanding the complexity of the system or intervention, and identifying key stakeholders.
- *Designing*: choosing appropriate evaluation methods, commissioning a complexity-appropriate evaluation, and involving stakeholders.
- *Embedding*: disseminating the results of evaluation and embedding the evaluation findings into policymaking.
- *Managing*: this phase does not have a chapter of its own, but is the overarching guidance which cuts across all three of the previously described sections.

---

[1] CECAN is a £3m research centre hosted by the University of Surrey. CECAN is working on methods and tools to improve the design and evaluation of policies related to the food, energy, water and environment 'nexus', and within areas where these issues interconnect in complex ways.

[2] Available at: http://sciencesearch.defra.gov.uk/Document.aspx?Document=14675_ComplexityEvaluationFramework.pdf

1.6   The CEF includes guidance, key questions and considerations, and links to suggested additional resources on particular evaluation methods. The CEF is comprises of a 30 page report detailing guidance, and a visual A3 summary poster.

Objectives and Approach of the Evaluation

1.7   The key objectives for this evaluation, as set out by Defra in the RfQ, are:

- To critically review the implementation of the CEF, identify any learning and areas for development, and provide practical advice and material for improvement. Questions of interest include:
    – How is the toolkit used in practice in Defra? Could changes to the CEF make it more efficient or effective to use?
    – Does the content add value? To what extent do users understand and apply the content? How could it better meet user needs and CEF objectives?
    – Is the toolkit, or aspects of it, more or less applicable in certain contexts? Such as at different phases of policy and evaluation or different types of users? How can this be further supported?
    – Are there clear examples of where the toolkit has improved the quality of evaluation design? How has the toolkit added value above and beyond what would have happened anyway? For example, the identification of risk, adoption of new methods.
    – Is there evidence of different, more flexible approaches to evaluation and policy making? For example, more adaptive and iterative practice, the inclusion (or discussion of the inclusion) of stakeholders.
    – Is there evidence of increased evaluation activity and complexity awareness? Are there early signs of embedding complexity thinking amongst users?
- Based on the review of the questions above, to develop a revised version of the CEF, if required, accompanied by user case studies (and any further resources), to support embedding use across Defra.

1.8   The evaluation of the CEF carried out included a combination of desk research, semi-structured interviews, observations and learning diaries.

Structure of this Report

1.9   The remainder of this report is structured as follows:

- Section 2 outlines the methodology used for the Evaluation;
- Section 3 outlines the scoping research undertaken and findings from the Steer-ED team's review of the CEF;
- Section 4 outlines findings from a desk-based review of Defra Invitation to Tender (ITT) documents;
- Section 5 outlines the findings from interviews and observations with CEF users;
- Section 6 outlines overall conclusions, changes made to the CEF with accompanying rationale, and suggested next steps.

    Details of specific changes made to the CEF are included in the Appendix.

2   Methodology

Approach

2.1   The evaluation sought to understand:

- The context for the work, including the objectives of the CEF, the development process, and the wider Government context.
- The degree to which complexity-aware approaches are being used in Defra evaluation commissions, both with and without reference to the CEF;
- How the CEF is being disseminated and used;
- Feedback on the CEF – including strengths and weaknesses, and user suggestions for further enhancements; and
- Early indications of whether the CEF is having an impact on evaluation practice in Defra, and if so, how.

2.2   The focus was therefore on undertaking a process evaluation, with some early examples of impact also being collected. The evaluation also sought evidence to contribute to revisions to the CEF, both directly (through collecting specific user feedback) and indirectly (via observation and consideration of how the CEF is being used).

2.3   The methods involved:

- Scoping interviews with Defra and CECAN staff who had been involved in the creation of the CEF, and with an evaluator from an Other Government Department (OGD), with related experience in developing evaluation guidelines for government;
- A desk-based review of the CEF by the Steer-ED team;
- A desk-based review of a range of recent relevant Defra Invitations to Tender (ITT) for evaluations;
- Observations of the CEF in use during one-to-many training events and one-to-one advisory sessions with Defra's Strategic Evaluation Team (SET) (who are involved with developing, promoting and using the CEF);
- Interviews with programme leads and analysts who had read, used or been informed about the CEF; and
- Learning diaries collected from SET, following one-to-one advisory sessions.

Scoping interviews

2.4   Semi-structured scoping interviews were conducted to help the Steer-ED team understand the context in which the CEF was developed, the methods used for developing the CEF, and the key objectives of the CEF. Interviews lasted up to an hour, and were conducted by one or two members of the Steer-ED team. Interview questions included:

- The interviewees' role in developing the CEF
- The objectives for the CEF
- What 'success' looks like for the CEF

- Barriers experienced in developing the CEF, and perceived barriers to achieving success
- Feedback received on the CEF during the development process.

2.5   The OGD evaluator, who had not been involved in the development of the CEF, was asked more general questions about the learnings from developing government evaluation guidance.

Desk-based review

Review of the CEF

2.6   Steer-ED undertook an independent review of the CEF. This looked at the 'framing' of the CEF, as an evaluation framework, a guidance document and a toolkit. It also reviewed each section of the CEF to test navigability, usability and intelligibility. It initially identified areas where modifications or edits may enhance the material.

2.7   The review also sought to identify the key components which a complexity-aware evaluation should consider, in order to inform a desk-based review of ITTs for evaluations in Defra's policy domains. The following provides a summary of specific recommendations from the CEF that were extracted from the CEF by Steer-ED, for use in the ITT review:

- Discussion of complex features of the system;
- Reference to systems mapping;
- Reference to logic modelling and/or Theory of Change;
- Review points and flexibility built-in to the evaluation design;
- Reference to stakeholders being engaged in the evaluation design;
- Methods cited are particularly suited to complexity-aware approaches – for example mixed methods, forward/backward looking methods, methods which capture both impact and descriptive or holistic features;
- Reference to the evaluation audience being primed to anticipate uncertainty; and
- Reference to multiple dissemination routes and/or plans to discuss early findings with key stakeholders.

2.8   Findings from Steer-ED's review of the CEF are reported in Section 3.

Review of Invitations to Tender (ITT) for evaluations

2.9   The Steer-ED team reviewed six recent, relevant ITT documents for planned or commissioned evaluations from Defra group, to scope and understand the treatment of complexity in the commissioning process. The review should not be considered a judgement of the quality of the ITTs considered or their ability to consider complexity. It is simply a means to identify signs, where they exist, that commissioners have followed recommendations from the CEF (if available) and, more broadly, the extent to which there is existing engagement with complexity aware approaches. While six is a relatively small sample, it was reasonable given the limited amount of time the CEF had been available for colleagues to use at the time of this evaluation.

2.10   For the purpose of the review, ITT documents were identified by SET to reflect a mix of policy domains and stages of evaluation. The authors of the ITTs had different levels of access to the CEF: two were in no way informed by the CEF, three had accessed guidance from SET (who are well-versed in the recommendations of the CEF), and one was produced with access to a full (draft, unpublished) version of the CEF as well as advice from SET. The purpose of the review was to identify the extent to which complexity-aware approaches were adopted with and without awareness of the CEF.

2.11   Findings from the review of ITTs are reported in Section 4.

Observation and interview methodology

2.12   The CEF has been designed with multiple types of user and uses across the broad spectrum of Defra's policy areas, and was anticipated to be disseminated using a number of different methods.

2.13   The CEF's audiences may be characterised as:

- **Commissioners/contract managers**, who will use the CEF to inform the design of evaluation invitation to tenders and the management of contractors;
- **Evaluators** (whether in-house or external contractors) who will use the CEF to conduct robust evaluations of complex interventions and/or interventions in complex systems; and
- **Policy leads** who may use the CEF to inform the design and/or commissioning of an intervention, and/or evaluation, as well as to interpret and respond to evaluation findings.

2.14   Three different dissemination techniques were observed by this evaluation:

- **One-to-one advisory sessions with SET**, which used the CEF to prompt discussion of complexity thinking in depth;
- **One-to-many training events** led by SET, which used the CEF as a teaching resource; and
- **Self-guided** use of the CEF.

2.15   Three types of data collection were conducted for this evaluation:

- **Observations** of one-to-one advisory sessions and one-to-many training events were conducted to gather information about how the CEF was used as a resource and approaches to delivering the advisory sessions.
- **Semi-structured interviews** (either by telephone or face-to-face) with users who had experienced each of the three different dissemination techniques (one-to-one advisory sessions, one-to-many training events, and self-guided use). These interviews were designed to gather feedback on: users' experience of applying the CEF; the lessons they had learned from it; their suggestions for changes or additional inclusions in the CEF; and finally, a description of how the recommendations from the CEF have contributed to their work, drawing evidence from a particular evaluation, if available.
- **Learning diaries** completed by the relevant SET colleagues after delivering one-to-one advisory sessions, to contribute their reflections on the CEF as a resource for facilitating discussion of complexity.

Interviewee selection

2.16   The Evaluation design aimed to achieve five interviews with users who had attended one-to-one advisory sessions and five interviews with self-guided users. This was achieved, and in fact exceeded for one-to-one advisory sessions[3]. The evaluation is felt to represent a sufficiently

---

[3] This is because some additional interviews had been built into the schedule in case of last-minute cancellations.

broad range of views, from a variety of different types of dissemination strategy, policy area, and stage of evaluation.

2.17   Interviewees were selected by SET colleagues from across the Department and its arms-lengths bodies on the basis of:

- Who was currently involved in an evaluation (whether that be in the initial scoping stage, the designing/procuring stage, or the final embedding stage), and there were clear indicators of complexity in the system or intervention being evaluated; and
- Ensuring a representative mix of different policy areas (summarised in Table 2-3), stages of evaluation, professional backgrounds and area of the department.

2.18   The above criteria were not applied to users who had attended a one-to-many training event – which were designed as a general introduction to complexity in evaluation. Instead, interviews were conducted with officials who had attended the session and volunteered to share their thoughts/reflections on the training and its relevance to their work.

Observation and Interview Process

For this evaluation, a total of 24 interviews and eight observations were conducted. These are summarised in summarised in Table 2-1 below, and further detail of each type of input is provided in

2.19   Table 2-2.

Table 2-1: Number of research inputs, by type

Type of briefing	Observations	Learning Diaries from SET officials	Phase 1 interviews (shortly after briefing)	Phase 2 interviews (longer term follow-up)
One-to-one advisory session participants	7	6	7	7
One-to-many training event participants[4]	1	n/a	1 group interview, 2 telephone interviews	n/a
Self-guided users	n/a	n/a	5	2[5]
TOTAL	**8**		**15**	**9**

---

[4] Two one-to-many training events were held during the course of this evaluation. One was an in-person introduction to evaluation, introducing the CEF, for officials working in the area of Future Farming, and the other was a Defra analyst webinar focussed exclusively on the CEF.

[5] Three of the self-guided users were not available for a second interview within the timeframe of the fieldwork.

2.20   Table 2-2 provides a summary of the approach. Table 2-3 lists the policy areas represented by the interviews with self-guided users and one-to-one advisory sessions.

Table 2-2: Outline of Evaluation Inputs

Type of briefing	Observation	Learning Diary	Phase 1 interviews	Phase 2 interviews
One-to-one advisory session participants	Session was observed in person by a researcher (or by telephone if the session took place as a teleconference).	Learning diaries were completed by the Defra evaluation official after the session, using a pre-agreed template.	Participants were interviewed in person immediately after the session[6]. Interview questions focussed on their feedback on the CEF, its relevance to their policy area, and how they might put the CEF recommendations into practice.	Participants were interviewed by telephone at least one month[7] after the first session. Questions focussed on understanding what information the user had retained, whether their views on the CEF had changed, and evidence of how the CEF was being used in practice.
One-to-many training event participants	Where possible, the training session was attended by a researcher either in person or by telephone.	n/a	Participants were interviewed either in person as a small group interview immediately following the session, or by telephone several days after the session had taken place. Interview questions focussed on the value of the CEF when used in this format, including what interviewees had	n/a

---

[6] In some cases, one-to-one advisory sessions were held with multiple individuals from the same policy area or team. In these cases, the team members were interviewed jointly for stage one and two interviews, according to individuals' preference/willingness to be involved.

[7] With a few exceptions, based on team availability. One interview took place 15 days after the first session, and three interviews took place slight less than one month later (between 26 to 30 days).

			learned from the session, what actions they had taken away, and whether they had any outstanding questions.	
Self-guided users	n/a	n/a	Users were interviewed either in person or by telephone (according to their preference) after having read the CEF. Questions sought to understand the experience of using the CEF as a stand-alone tool, feedback on the CEF, actions taken away, and outstanding questions.	Users were interviewed by telephone at least one month after their first interview. Questions focussed on understanding what information the user had retained, whether their views on the CEF had changed, and evidence of how the CEF was being used in practice.

Source: Steer-ED, 2020

Table 2-3: Details of policy areas represented by interviews conducted

Policy Area
Future Farming & Countryside
Marine & Fisheries
Natural Environment
Environmental Land Management
Chief Science Advisor's Office
Environmental Quality (including Air Quality, Resources & Waste and Chemicals & Pesticides)
EU Exit and International Trade

Source: Steer-ED, 2020

2.21   To ensure a full and frank discussion with interviewees, it was agreed that their responses would be reported on a non-attributable basis. Views are therefore summarised in general terms, without specifying the sources of comments.

2.22   Findings from observations and interviews are reported in Section 5.

Limitations

2.23   The evaluation was undertaken during the early phase of CEF roll-out. Limitations of the study are that:

- A relatively small number of CEF users were available to be interviewed (since the CEF was not yet in widespread use across the department);
- Interviewees had relatively limited experience using the CEF (since they had been using it for a few months only). This limited the depth and breadth of the interviews and observations conducted; and
- There was no opportunity to examine long-term usage patterns or the extent to which interest in the CEF changes over time.

3 Scoping review of the CEF

3.1 Steer-ED conducted an independent review of the CEF and undertook scoping interviews with Defra and CECAN staff who had been involved in the creation of the CEF, and with an OGD evaluator. This review, along with the scoping research undertaken, informed the lines of enquiry taken for the rest of the Evaluation. It also fed into decisions made by Steer-ED about modifications to the CEF (which are described in detail in Section 6).

Steer-ED's Review of the CEF

3.2 An initial review, prior to speaking to CEF developers or users, brought up the following points from the Steer-ED team:

- The team agreed that the CEF provides a helpful summary of the challenge of evaluation in the face of complexity.
- The advice goes into the right level of detail, offering clear guidance that would be accessible to a wide range of different users.
- The questions raised are clear, helpful and likely to spark new ideas or approaches to evaluation, and the table of key questions is particularly useful.
- In comparison to other comparable resources (for example the Impact Evaluation Framework for Regional Development Agencies)[8], the CEF seems to focus more heavily on key questions and 'things to think about' rather than specific guidance. The team felt this could leave users confused or unsatisfied.
- The team questioned the appropriateness of the term 'toolkit', given the above.
- The team thought the 'Visual Guide to Complexity' was informative and easy to follow, and introduced key terms that could help colleagues discuss complexity and how to study it.
- The team found the A3 poster difficult to read.
- The team felt it would be difficult to read the CEF as a printed document, given the number of hyperlinks.

Scoping stage

3.3 The direction for the Evaluation was guided by the Scoping stage. Main findings from the scoping interviews were that:

- There is already a background understanding of complexity within Defra, and expertise within the department. Interviewees explained the existing relationship with CECAN, who have helped to elevate the role of complexity thinking within the department through seminars and the provision of resources.

---

[8] **https://www.sqw.co.uk/files/4813/8712/1417/149.pdf**

- The CEF is intended to be accessible to a range of different users, from different disciplines and with different levels of experience with complexity-appropriate evaluation. When asked about the development of the CEF, interviewees stressed the importance of creating a document which is concise, easy to use and avoids the use of jargon or overly technical language. This also determined the concept of the CEF as a 'signposting resource' rather than an all-encompassing guidance document.

- It was considered to be important to provide a document tailored to the Defra audience rather than generic complexity guidance. This point was made by Defra staff interviewed, and also echoed by the OGD evaluator, who explained that guidance documents are usually more effective when tailored to the intended audience. This is because readers may dismiss examples drawn from other policy domains as being non-transferable or irrelevant to their own policy area.

- The level of use of the CEF, and the extent to which it informs evaluations, will define its success. Steer-ED asked 'what does success look like for the CEF', and the responses included – that the CEF is used frequently, brought along to meetings, shared amongst staff, and used to brief new members of staff. In addition, interviewees explained that they hoped the CEF would result in enhancement and more wide-reaching use of complexity-aware methods throughout the policy evaluation life-cycle in Defra.

- Steer-ED asked what barriers had been experienced during the CEF development process. Responses were mostly around ensuring the tone and content were correctly pitched, such that the output document was useful without being prescriptive; and thought-provoking without being confusing.

- Interviewees were also asked what barriers they anticipated might prevent widespread adoption of the CEF. They mentioned the importance of the right dissemination strategy, such that the CEF receives sufficient attention amongst the many other guidance documents that exist. Interviewees also expressed some concern about how users might respond to the CEF, and whether some may find it off-putting or unhelpful (for example by over-complicating issues or failing to address some of the more intractable issues users are grappling with).

Summary

3.4 The information gathered as part of the review of the CEF and scoping interviews was used to provide Steer-ED with context and background information to inform the evaluation approach. Topic guides for interviews and observations were developed to take into account what had been learned at this stage, including the aims of the CEF, decisions taken during the development of the CEF, and any concerns or potential barriers identified.

4 Review of Invitations to Tender

4.1 As noted in Section 2, six recent ITT documents were identified for review by SET and Defra colleagues to reflect a mix of policy domains and stages of evaluation. The purpose of the review was to understand and identify the extent to which complexity-aware approaches were adopted with and without awareness of the CEF. For the purposes of this review, the evaluation ITTs examined will be referred to as documents A to F.

Headline analysis

4.2 Table 4-1 provides a summary of the key findings from the review. It sets out the eight key recommendations Steer-ED identified as being made by the CEF for those designing and managing complexity aware evaluations. Wherever evidence of a CEF recommendation was found in an ITT, the relevant ITT has been noted using a unique letter (A to F).

4.3 As discussed previously, those working on the ITTs had varied levels of engagement with the CEF. Results are therefore presented in three columns: from those which were in no way informed by the CEF (documents A and B), to those which were to some extent informed by the CEF (documents C, D and E) and one which was developed with access to a full (unpublished) version of the CEF (document F).

4.4 The colour coding in the table indicates the proportion of documents for which evidence in line with a key recommendation was found. Red means no evidence was found; Amber means 50% or fewer contained evidence of the recommendation; and Green means more than 50% of the tender documents contained evidence of the recommendation. Note that no evidence of complexity-aware approaches was found in ITT document B, and so it is absent from Table 4-1.

Key Findings

4.5 Based on this selection of documents, engagement with the CEF does seem to be related to greater adoption of the CEF's key recommendations, as indicated in Table 4-1. There are positive indications that, even at this early stage, users have taken on board some of the lessons from the CEF and started to put them into practice. The review found that users with access to the CEF were more likely to refer to systems mapping and participatory approaches to evaluation design than those without.

4.6 The review found that users from almost all groups, even those who did not have access to the CEF or only had 'light-touch' access, were able to identify and communicate some of the complex features of the evaluation system. This demonstrates the existing complexity knowledge in the department. In some cases it may have been deemed unhelpful or excessive to include details about the complexity of the system in an ITT. We therefore do not infer that absence of complexity in an ITT means a lack of knowledge of complexity.

4.7 Finally, there was little evidence of the CEF's recommendations in relation to the *embedding* of evaluation findings (that is, the plan for disseminating and using evaluation findings), even

amongst those ITT's developed with reference to the CEF. However, note that this may be because many of the 'embedding' activities occur after the design and delivery stage, and so would not necessarily be referred to (or even known) at the evaluation tendering stage.

Table 4-1: Summary of findings from review of six tender documents

Level of engagement with the CEF	Not informed by CEF (n=2)	Informed by CEF (n=3)	Used CEF (n=1)	Total
Document references of ITTs examined	A,B	C,D,E	F	**6**
Discussion of complex features of the system	A	C, D, E	F	**5**
Reference to systems mapping		E	F	**2**
Reference to logic modelling and/or Theory of Change	A	C, D	F	**4**
Review points and flexibility built-in to the evaluation design	A	C, E	F	**4**
Reference to stakeholders being engaged in the evaluation design		C	F	**2**
Methods cited are particularly suited to complexity-aware approaches – for example mixed methods, forwards/backwards looking methods, methods which capture both impact and descriptive or holistic features.	A	C, D	F	**4**
Reference to the evaluation audience being primed to anticipate uncertainty				**0**
Reference to multiple dissemination routes and/or plans to discuss early findings with key stakeholders	A			**1**

Note: The leftmost column of the table is labelled vertically "Key recommendations from the CEF".

Source: Steer-ED analysis of Defra ITT documents, 2019

4.8   Table 4-2 summarises the evidence of 'how' recommendations from the CEF had been adopted or utilised in the six ITT documents reviewed, for example, specific reference to 'rapid change' and 'uncertainty', stakeholder engagement to reflect the complexity of actors in the system and so on.

Table 4-2: Evidence found in the ITTs reviewed that key recommendations from the CEF have been adopted

CEF recommendations	Evidence in ITTs that the recommendation had been adopted
Discussion of complex features of the system	• Recognition that some impacts may be felt by individuals who were not the direct recipients of the intervention • Reference to multiple different agencies/bodies involved, operating at different levels (for example national, regional, local and individual) • Reference to rapid change within the system, and uncertainty of when/where impacts will appear
Reference to systems mapping	Mention of systems mapping as a tool for understanding complex policy areas.
Reference to logic modelling and/or Theory of Change	Requests or requirements for suppliers to develop logic models or theories of changes as part of the evaluation.
Review points and flexibility built-in to the evaluation design	• A feasibility/scoping stage built into study design (which helps to identify and manage risks associated with matching evaluation methodologies, research questions and budgets for complexity-aware evaluations). • Built-in evaluation points where the study steering group will decide whether or not to proceed. • Recognition that circumstances may change, and tenderers asked to demonstrate how their evaluation approach will be adapted to changing circumstances.
Reference to stakeholders being engaged in the evaluation design	Expert knowledge of stakeholders acknowledged. Tenderers expected to draw upon stakeholders to help develop systems maps, provide data, and be involved in the scoping and design of the study.
Methods cited are particularly suited to complexity-aware approaches – for example mixed methods, forwards/backwards looking methods, methods which capture both impact and descriptive or holistic features.	• Recommendation that a range of different methods (empirical, theory-based, case studies and qualitative research) be used in conjunction. • Proposed developmental evaluation approach (a method particularly appropriate to evaluation in the presence of complexity). • Expectation that the evaluation will be an 'ongoing process of learning' rather than a backward-looking assessment.
Reference to the evaluation audience being primed to anticipate uncertainty	No evidence found.

Reference to multiple dissemination routes and/or plans to discuss early findings with key stakeholders	Reference to adaptive policy making, including a plan to use interim evaluation outputs to inform policy decisions. Tenders invited to include activities which will support adaptive policy making.

Source: Steer-ED analysis of Defra ITT documents, 2019

5 Findings from observing the CEF in use

5.1 This Section presents findings from observations and interviews conducted with Defra group staff, almost all of whom were actively using the CEF in developing evaluations[9]. This Section presents: evidence of how the CEF is being used and the extent to which the recommendations of the CEF are being adopted; feedback from interviewees on the experience of using the CEF and suggestions for additional material which could be included; and finally, feedback on the dissemination strategy of the CEF.

5.2 The following Section (Section 6, Conclusions and Recommendations) shows how the feedback received was taken forward to make changes to the CEF.

Evidence of how the CEF is being used

5.3 As noted previously, the CEF is in its early stages of use, which means that to date there are a limited number of examples of its application and impact. Table 5-1 provides initial findings in relation to application and impact of the CEF. Overall, it shows that interviewees who have applied the CEF to a specific evaluation have begun to implement some of its recommendations, in the following ways:

- By being able to better understand and more clearly articulate the complex features of the system they are working with.
- By taking a more holistic approach to the evaluation, and moving away from a desire for a simplistic or linear causal understanding.
- By adopting a more flexible, stakeholder-led evaluation design.
- By considering the implications for reviewing and disseminating the results of the evaluation, given the complexity of the subject.

5.4 The findings presented in the table below should be treated as a snapshot of early impacts – further impacts are likely to arise over time; equally short-term impacts may fade without steps to maintain momentum in embedding complexity thinking in Defra's approach to evaluation.

Table 5-1: Examples of how the CEF is changing thinking/actions

Stage	Theme	Comments
Understanding	Helping people to think more broadly and holistically about an evaluation – its	Interviewees commented that the CEF: • Helped them to focus on the 'bigger picture'. • Helped them to think more holistically about the evaluation, with all of its various interlinking parts.

---

[9] The exception is those interviewees who had attended one-to-many training events.

Stage	Theme	Comments
	context, outputs and stakeholders.	• Had resulted in them considering additional ways to involve stakeholders in evaluation design, for example by considering participatory approaches. *"[Reading the CEF] has made me think a lot more holistically about complex evaluation. It's given me a structure to think about it. The evaluation I'm working on is so complex, with so many different strands, all wriggling out in different directions. So it helps me to have a framework if you like, a way of thinking of it all as one piece"*
Designing	Giving people confidence and tools for dealing with a 'messy' evaluation	Interviewees commented that the CEF: • Gives reassurance, helping them to acknowledge and manage complexity in an evaluation. • Provides a useful language to communicate complexity, both internally and externally. Many found the descriptions of characteristics and behaviours of complex systems provided helpful language that they hadn't previously used to describe the system. • Has helped them to move away from simple linear 'logic-chain' approaches and towards a more integrated, iterative approach. • Resulted in them putting greater emphasis on Theory of Change and systems mapping as tools. Some, on reading the CEF's advice, thought that they could have made greater use of systems mapping in previous evaluations. • Had resulted in greater consideration of adaptive policy making approaches and/or methods which deliver interim results to support the policy process. • Had given them additional knowledge to improve the management of sub-contractors working on complex evaluations. • Had encouraged consideration of a more flexible evaluation design. • Had encouraged them to consider more regular review of monitoring data, and to think of the evaluation in terms of a continuous feedback loop rather than a 'one-shot' event. *"It [reading the CEF] has made me really think about it* [evaluation] *not being one-off, being a continuous feedback loop. For some of the*

Stage	Theme	Comments
		interventions, it's made me think about <u>when</u> we should be looking for results."
Embedding	Encouraging those managing evaluation to communicate findings in a more nuanced way	Interviewees commented that the CEF: • Gave helpful advice about how to manage expectations of stakeholders, both internally and externally. • Gave a useful checklist of advice/issues to raise when disseminating findings from an evaluation. • Recommended positive, clear language about how to explain the limitations of results, helping to avoid what could be perceived as negative language about the limitations of results. *"It can sometimes come across like you're trying to make excuses [when putting caveats around the difficulties of evaluation in the presence of complexity], but it didn't have that tone at all."*

Source: Steer-ED, 2020

Feedback on interviewees' experience of using the CEF

5.6 Table 5-2 provides a summary of all feedback received from interviewees on the experience of using the CEF. Interviewees were asked to comment on the sections of the CEF they found most and least helpful, and also invited to suggest additional content which they would find helpful. Responses were concentrated around certain sections of the CEF, such as the 'Visual Guide to Complexity' for example – which was frequently discussed by interviewees[10]. Many interviewees also went further than citing only the single most and least helpful parts of the CEF, commenting too on other sections which they liked, found useful, or did not like/find useful. All responses to this line of questioning are reported in Table 5-2 below.

5.7 The results outlined below reflect the most commonly occurring themes or viewpoints collected during the evaluation. The views of individual interviewees are presented only where they align with a comment or broader theme also raised by other interviewees.

5.8 The feedback reflects the diversity of the interviewees consulted – who ranged from experienced evaluators to those more experienced in other domains, and from scientists who were familiar with complexity thinking to evaluators who were newer to complexity theory. It indicates that on the whole the CEF:

• Provides a useful overview and introduction to complexity thinking. The majority of users stated that they found it useful, would refer to it again, and would recommend it to colleagues;

---

[10] In part this is because the Visual Guide to Complexity was used as a prompt to begin a discussion about the complex features of the system in one-to-one advisory sessions.

- Acts as an aide-memoire to those already experienced with complexity-aware evaluation, and an entry-point for those users who are new to complexity thinking;
- Strikes a balance between the:
  - Need for technical precision with approachability;
  - Provision of practical guidance in sufficient depth to understand the issues to be addressed, with appropriate signposting to other sources where more in-depth coverage is required.

5.9   In making amendments to the CEF, users considered that the CEF may benefit from:

- Additional introductory material that puts complexity-thinking and evaluation in context and helps users to understand how to get the most from the document;
- Editing in some places to provide more straightforward language where terminology dominates;
- More prominent signposting to relevant evaluation methodologies without overburdening the CEF with material that is readily available elsewhere;
- Revisions to the formatting of the A3 summary, to enhance clarity and readability.

Table 5-2: Feedback on the CEF[11]

Theme	Summary
Overarching feedback	
Comments on overall tone, language and style	The following comments were made by interviewees: **Strengths** • The length is appropriate • The writing style is clear, succinct and not overly academic **Challenges** • It can be repetitive • It uses too much technical language or 'jargon'
Comments on content	The following comments were made by interviewees: **Strengths** • It is comprehensive, successfully pulling together key materials, signposting and considerations in one document. **Challenges** • The weighting being towards questions rather than answers (or prescriptions) is unhelpful • The overlap with best-practice evaluation guidance from outside the complexity domain is confusing • As a stand-alone document, there is insufficient guidance to allow a user with little experience of complexity-aware evaluation to arrive at a suitable evaluation design.

---

[11] Note all page references refer to the first version of the CEF, rather than the revised version which was produced as a product of this evaluation.

Theme	Summary
	Most interviewees cited either the *Designing* or *Embedding* sections (or both) as being the most useful parts of the CEF. Little comment was made on the *Understanding* section, which was generally considered to be mostly background information. Although interviewees mostly found the *Designing* section useful, some felt that it was insufficiently prescriptive, focussing too much on what *could* be done rather than what *should* be done. Some interviewees were also left feeling confused as to which method they should select. For the *Embedding* section, there was general agreement that it provided helpful content, with some interviewees stating that they would likely return to this section at a later stage in the evaluation process. Some interviewees raised concerns around the feasibility of the recommendations, for example that they might be practically difficult to implement if the evaluation had already been designed, without prior reference to the CEF.
Feedback on specific sections/content	
A visual guide to understanding complexity for Defra (pg 7-8, and Annex I)	Most interviewees referred to this content and used it as part of their identification of complexity in the system. Some were positive, and others raised challenges. **Strengths** • Interviewees found it a helpful resource, which assisted with understanding or describing complex systems. "[The] *complexity examples were very good – definitely accessible and not too social science heavy"* **Challenges** • Some found the examples to be too technical, and two interviewees found the examples too natural-world focused. • There was some uncertainty around the structure – overlaps between some of the definitions, and confusion around the division of examples between the annex and body of the text. • Finally, more than one user questioned what the intended use for the characteristics section should be, beyond simply checking which characteristics apply to the policy area of interest. *"I have mixed feelings about that* [the characteristics of a complex system]. *They're useful ways to think about the system, but on the other hand some of it is quite techy. I'm not sure how relevant it* [the section] *is really."*
Table of key questions (pg 14-15)	The majority of interviewees made use of the tables of key questions, either in the one-to-one advisory session observed, or on their own while working through the CEF. Many interviewees

Theme	Summary
	commented that they found the questions to be extremely useful prompts.
	Not all interviewees realised that the A3 summary document reproduced these key questions, due to the formatting differences.
References to RCTs (pg 21, 22)	Several interviewees raised questions or issues around the treatment of RCTs in the CEF. One found the user quote (page 21) to be unclear, another found the discussion of RCTs (page 22) to give insufficient guidance.
A3 Summary	Many interviewees commented on the A3 summary document, which elicited strong opinions. **Strengths** Those who liked the A3 thought it was extremely useful to have everything on one page, with one reporting that they had attached it to their wall. Several interviewees commented that while they had initially found it off-putting, once they had read the full guidance document, they began to find it useful. **Challenges** • There was an apparent issue amongst self-guided users of the CEF, who were not necessarily aware of the A3 as a separate file available for download. • Some interviewees found the text is too small for some users to read • Several commented that they found it difficult to know how to read it – with no clear top to bottom/left to right sequence for the reader. *"I needed to put a decent pair of glasses on! It was helpful in covering all the aspects, but it took a while to work out how it all fits together. Whilst I appreciate the integrated and feedback loop type approach, it would be useful to have a beginning and an end."*
Suggestions for Additional Material	
Explanation of the benefits of evaluation – why evaluate in the first place, and how to make the case for evaluation	Two interviewees commented that there was insufficient guidance on the purpose and value of evaluation in general, which would help to set the context for complexity-appropriate evaluation.
Explain how the CEF sits within the wider context of good-practice evaluation	Several interviewees raised questions around the 'uniqueness' of complexity – they wondered to what extent the guidance is simply a repeat of existing good-practice evaluation.

Theme	Summary
Additional information about choosing appropriate methods	On the subject of methods, interviewees raised the issues of: • Too few examples of methods in use – for example one user asked for an example ITT to be included in the CEF. • Not enough explanation/direction of how to choose an appropriate method. • Much of the advice concentrates on very complex or technical approaches, and may mislead readers into thinking 'simple' methods are not appropriate. • Insufficient discussion of how to 'flex' methods or approaches according to the study budget or other resource limitations.
Other requests for additional content	Interviewees made the following other requests for additional content: • Case studies or vignettes showing what complexity-aware evaluation looks like • Advice on monitoring in the presence of complexity • An executive summary • An explanation of how evaluation quality scales treat complexity-aware evaluations • Additional guidance on the role of Randomised Control Trials (RCTs).

Source: Steer-ED, 2020

Dissemination of the CEF

5.13   Interviewees were also asked to provide feedback on the how the CEF had been shared within Defra. Many had first come across the CEF through the Defra Evaluation Community of Practice[12]. For those who had attended a one-to-one advisory session with SET, there was general agreement that the briefing gave a useful introduction to the tool and the availability of internal expertise to assist with complexity-aware approaches.

5.14   Self-guided users in general seemed to gain useful advice from the CEF, and did not differ significantly from one-to-one advisory users in terms of their comments or their ability to put the CEF advice into practice. This is a positive sign that the CEF may function effectively as a stand-alone resource with some audiences. Several interviewees however did seem at a loss as to how to handle questions they had after reading the CEF, and it did not always occur to them where they might find expert support on this subject. It therefore would seem worthwhile including signposting to SET or Defra analytical resources as part of the CEF dissemination strategy.

5.15   Based on the observations, feedback and learning diaries collected regarding one-to-one advisory sessions, offered below are a few points of best practice for the delivery of future advisory sessions by SET, or other Defra evaluation specialists. Various different delivery/dissemination styles were observed, and the points below are suggestions based on what seemed to work best (based on Steer-ED's reflections and the views heard from interviewees):

---

[12] This is an open group for Defra group officials to share ideas on evaluation best practice.

- Taking time to introduce and 'advertise' the CEF, including sending users a link to it in advance of the session, and potentially bringing a copy of the document along to the meeting;
- Agreeing with users before the session which stage of evaluation they are contemplating (*Understanding*, *Designing* or *Embedding*), and then focussing the session on the section(s) of specific interest;
- Bringing paper copies of resources seems to help guide the session. Suggested resources are: a copy of the CEF guidance document, to familiarise users with the resource; a copy of the A3 summary; and/or a copy of the key questions tables (page 14-15) as well as the visual guide to complexity. When delivering the session via teleconference it would be worth considering how to provide suitable resources in a way that all attendees can follow, for example using screen-sharing functionality.

5.16   On the subject of one-to-many training events, the evidence collected indicated that attendees were positive about the content covered, which served as a good introduction to the concepts of complexity. However, they did not seem to have engaged with the CEF to the same degree as those who had attended a one-to-one advisory session, probably because they had not gone through the same process of applying it to a specific evaluation within their policy area. These training events should therefore not be seen as a substitute for one-to-one advisory session, but instead as more of an exercise in awareness raising.

6   Conclusions & Recommendations

6.1   This section brings together findings from the desk-based review and observations of the CEF in use, and makes recommendations. It has three sub-sections:

- Overall findings in relation to the CEF, its rollout and its application to date;
- A summary of findings which indicate (either directly or indirectly) suggested modifications to the CEF, and Steer ED's response to each of these; and
- Recommended next steps in continuing to develop the use of CEF within Defra.

Overall findings

6.2   This evaluation collected information on the use of the CEF at a range of different stages of application in Defra, and feedback from users on their experience interacting with the resource. The evaluation took place with a small number of users, at an early phase of CEF use - the conclusions that follow should therefore be considered within this context. The toolkit is being used as both a stand-alone ('self-guided') resource and as a resource to structure and inform advisory sessions delivered by Defra evaluation specialists. In both cases, those interacting with the CEF were able to derive useful learning and ideas on handling complexity evaluation from what they read, heard or discussed.

- It has also been successfully presented at one-to-many training events, which has engaged the interest of attendees and helped to introduce the concepts of complexity-aware evaluation.
- Although still in early stages of use, there is already some evidence that the CEF is changing behaviour – for example by encouraging a more holistic or nuanced approach to evaluation, the adoption of complexity-appropriate methods, or more stakeholder-driven approaches.
- There are some clear examples of complexity-aware evaluation designs being used in Defra, and these were evident from the review of ITTs. It is difficult to draw causality back to the CEF however – as the ITTs do not specifically refer to the CEF as the source of guidance on complexity-aware approaches.
- While the CEF was in relatively early release phase at the time of the evaluation, some interviewees indicated that they had absorbed the guidance, to some extent at least, and were likely make use of a complexity-aware approach in future. Other interviewees, who were already well-grounded in complexity-aware approaches, reported that the CEF provides reassurance and a comprehensive reference tool.

6.3   Overall, there is no question that the main content of the CEF, in terms of the key considerations and questions, is proving useful and is already contributing to adoption of some complexity-aware evaluation practices in Defra.

6.4   Based on Steer-ED's review, and the feedback received from interviewees, there are some opportunities to augment the value of the CEF – by improving the clarity in some areas,

offering additional introductory text, and giving thought to the framing of the CEF. These changes are discussed in the following Recommendations section.

Recommendations for amendments to the CEF

6.5   Table 6-1 contains a summary of opportunities for improvements to the CEF which have been identified through the course of this Evaluation. These are based on interviewees' feedback, Steer-ED's review of the CEF and observations of how the CEF is being used. The points are accompanied by a response (in terms of modifications to the CEF, or otherwise) and a rationale for the response (including a discussion of why some changes were taken forward, and others were not).

6.6   For details of the specific text inserted into the CEF, refer to Appendix A. Note that minor changes, such as correcting typos, or updates to ensure consistency (such as numbering or cross-references) have not been described. Other changes made to the CEF which have not been presented in Appendix A are:

- The inclusion of a set of 'user journey' case studies, collected during this evaluation.
- Redesign of the A3 summary poster.
- Amendments to the Visual Guide to Complexity, to include some additional (non-natural world focussed) examples.

6.7   Steer-ED analysed all interview feedback and made recommendations for changes based on key themes. In some instances, changes were suggested based on feedback from a small number of interviewees, where it was corroborated by other evidence and felt that these changes would optimise the use of the CEF for all users.

Recommended next steps in continuing to develop the use of the CEF within Defra

6.8   This evaluation was conducted shortly after the first publication of the CEF. It has collected early evidence of how the guidance is being used and applied in Defra. Given the timing, the findings of this study will have been limited as many interviewees had only had relatively short exposure to the CEF, and no users had yet experienced using it from the first 'understanding' stages of an evaluation all the way to the final 'embedding' stages. Steer-ED therefore recommend a follow-up review should be carried out in approximately one to two years' time. This would help to:

- Provide a fuller picture of how the CEF is used, from start to finish of an evaluation;
- Test whether the CEF continues to be of relevance, given changing political contexts and government priorities; and
- Test whether the CEF continues to be of relevance once the revised Magenta Book annex on complexity is embedded, and whether it needs to be repositioned in any way.

6.9   Steer-ED also recommend that, given the quantity of hyperlinks contained in the document, a regular update process should be undertaken to ensure these links remain active.

Table 6-1: Feedback on the CEF/opportunities for improvements, and Steer ED's response

Finding	Steer ED's Response	Rationale
Navigability of the document On discussion with interviewees, Steer-ED observed that many users were not getting the most out of the document, because they had not followed the hyperlinks, or were not aware that the A3 was available as a separate file.	**Addition** Inclusion of an insert in the CEF report titled 'How to use this document', which recommends reading the document on-screen, and which also references the A3 summary. Inclusion of URLs in the Resources appendix.	Steer-ED felt these additions would help users to get the most out of the document.
Interviewees suggested it would be useful to include an explanation of the **benefits of evaluation** also of how the CEF sits within the **wider context of good-practice evaluation.**	**Addition** Inclusion of additional text in the Introduction of the CEF which explains the benefits of evaluation, the resources available for good-practice evaluation, and the role of the CEF, including citing the Magenta Book.	Additional text clarifies the questions/uncertainties that interviewees raised when reviewing the content of this section.
Interviewees asked for additional information about **methods –** specifically, how to choose a suitable research method, whether 'simple' methods are appropriate for complex evaluation, and what role RCTs can play in complex evaluation.	**Addition** • Addition of a paragraph which discusses the pros and cons of different tools, and the role of different experimental methods. • Addition of text to describe the 'Choosing Appropriate Evaluation Methods' tool[13] in more detail.	Steer-ED felt that additional text elaborating on the pros and cos of different methods would help to reduce confusion about whether certain methods were recommended/advised against for complexity-aware evaluations. Much of the content users requested regarding methods (in particular, how to choose between methods) is in fact already signposted in the CEF. The 'Choosing Appropriate Evaluation Methods' tool (Befani, 2016) is a particularly

---

[13] Befani, B. (2016), *Choosing appropriate evaluation methods* [tool], London: Bond. Online at: **https://www.bond.org.uk/resources/evaluation-methods-tool**

		useful resource for this. Steer-ED concluded that interviewees were missing out on valuable content because they had not followed the links in the CEF. Steer-ED therefore chose to provide additional information about this toolkit to encourage users to visit the webpage.
The '**visual guide to understanding complexity**' was found to be useful by most interviewees, however some felt it was too technical or too natural-world focussed. There was also some uncertainty around the structure and purpose of the section.	**Addition** Inclusion of additional text which: • Explains the rationale for dividing the examples between the annex and the main body of the text • Reassures the reader that there may be some overlap between examples • Explains the purpose of the section. Inclusion of some examples which are not drawn from the natural-world.	Additional text clarifies the questions/uncertainties that interviewees raised when reviewing the content of this section.
Table of **key questions** (pg 14-15) – many interviewees cited this as the most helpful part of the CEF.	**Addition/amendment** • Inclusion of an additional Annex, duplicating this content, for ease of printing. • Reformat the A3 so that it is clearer that it contains these key questions.	Steer-ED agreed with interviewees that this table was one of the most useful pieces of content in the CEF. The response was therefore designed to raise the prominence of these tables and make them easier to use as a stand-alone resource.
Interviewees requested for additional guidance on the role of **RCTs**, and amendments to the discussion of RCTs.	**No change**	Steer-ED felt it would not be appropriate to provide significant additional material on one method alone. Instead, Steer-ED choose to give additional, more general, advice and signposting about methods.

The **A3 summary** elicited very mixed responses. Some found it extremely useful; but many suggested the text was too dense and difficult to navigate. This was also the view of the Steer-ED team.	**Amendment** • Amendments to the A3 summary to improve the readability and guide the user through the page. • Reformat the questions on the A3 summary so that they are clearly recognisable as the tables from the main document. • Remove the reproduction of the A3 summary in the main document, since the improved navigability means this is no longer necessary.	Steer-ED felt that improvements could be made that would enhance the A3 for all users - both those who already found it useful, and those who found it difficult to read in its current format.
Request for inclusion of **case studies** or vignettes showing what complexity-aware evaluation looks like.	**Addition** Inclusion of a separate annex of case studies.	The majority interviewees made this request, so Steer-ED felt it was an important addition to the CEF.
Request for inclusion of advice on **monitoring** in the presence of complexity.	**Addition** Inclusion of additional text on the correct treatment of/importance of monitoring data in the presence of complexity.	Steer-ED agreed that this would make a valuable addition.
Request for an **executive summary**.	**Amendment** Improvements to the A3 summary, as described above, so that it effectively functions as an executive summary.	It was requested that Steer-ED avoid creating an additional summary document, and instead focus on improvements to the A3, allowing it to function as an Executive Summary.
Request for an explanation of how evaluation **quality scales** treat complexity-aware evaluations	**No change**	It was requested that Steer-ED do not include reference to evaluation quality scales since these can have differing interpretations, and may not sufficiently take into account complexity.
Interviewees asked for more discussion of how to respond to resource limitations and advice on	**Addition**	Steer-ED agreed that this would make a valuable addition.

how to approach complexity-aware evaluation with **proportionality**.	Addition of text on 'proportionate treatment of complexity', which will address relative costs and benefits of different research methods relative to the scale of programme and/or its likely impact.	
Interviewees asked for **practical advice on how to commission**, procure and manage a complexity-aware evaluation.	**Addition** Inclusion of a text box containing considerations for commissioning, drawn from best practice advice and observations from the desk research and interviews.	Steer-ED agreed that this would make a valuable addition.
Some interviewees were unclear about the difference between a **complex intervention and a complex policy area**. Some asked whether complexity is relevant to all policy areas, and whether they could see examples of 'non-complex' evaluations.	**Addition** Inclusion of explanatory text which clarifies this issue.	Steer-ED also felt this was unclear in the CEF and thought it would be worthwhile including additional text.

7   Appendix

Amendments or additions to the text in the Complexity Evaluation Framework are detailed in the tables below.

Section 0: 'How to use this Document'

This document is divided into two parts:
Part 1: What is complexity and what does it mean for Defra? (Chapter 1) explores what we mean when we talk about complexity, makes the case for why evaluation is crucial when working with complex systems, and introduces the core principles of complexity-appropriate evaluation.

Part 2: The Complexity Evaluation Framework (Chapters 2 to 5) sets out a practical framework of considerations and guidance for those designing, managing and embedding evaluations.

The document is intended to be used as a reference guide. The reader is encouraged to begin with Part 1, and then to concentrate on the sections of Part 2 most relevant to the evaluation element they are working on. A set of user case studies are provided in Annex VI which describe how others have used the CEF to work with complexity-appropriate evaluations, at various stages of development.

Hyperlinks throughout this document give signposts to useful additional resources. To get the most out of the document, reading on-screen/digitally is recommended. A printable 'quick-reference guide' is also available in A3 poster format.

Section 0: 'What is the Complexity Evaluation Framework?'

The Complexity Evaluation Framework (CEF) is designed to assist policy makers and analysts to design and deliver effective evaluations under complex circumstances.

A system or process that is complex is made up of many diverse components that interact in nonlinear ways and may also adapt or change over time. This can lead to unpredictable behaviour and unexpected outcomes. The domains that Defra deals with are complex and can involve working with complex environmental and social/economic systems, often at the same time.

The purpose of the CEF is to guide the scoping, commissioning, management and delivery of evaluations in the presence of complexity. It provides guidance on complexity characteristics and a framework of considerations to inform conversations between policy leads, commissioners of evaluation, and evaluation contractors. Its aim is to ensure that complexity is recognised and complexity thinking is embedded into evaluation design and delivery to ensure approaches are robust. The CEF is intended to support evaluation

across the policy cycle, from as early in the policy cycle as possible. It could also be used outside of evaluation, including in policy scoping and option appraisal.

This is the second version of the Complexity Evaluation Framework, which has been updated following an independent evaluation of the CEF in its early implementation. It was originally developed for Defra by the Centre for Complexity Across the Nexus (CECAN), through the engagement of Defra staff and use of academic literature. The first published version was evaluated by an independent evaluator, Steer Economic Development (Steer-ED), who made recommendations to help maximise the value of the CEF and collected case studies of early uses of the CEF (see Annex VI). This second version reflects changes made in response to those recommendations. Full details of Steer-ED's evaluation can be found in their published Evaluation report.

Section 0: 'Why Evaluate?'

Evaluation is an integral part of the government policy making ROAMEF cycle (which stands for Rationale, Objectives, Appraisal, Monitoring, Evaluation, and Feedback), and is formalised in the HM Treasury Magenta Book and Green Book Evaluation. When performed well, evaluation can give the policy maker valuable, objective, insights as to:

- The impact or effectiveness of a policy, including both anticipated and unanticipated effects, providing or enhancing the evidence-base that supports policy decisions;
- An understanding of how the intervention worked;
- How successfully it was delivered;
- Whether it generated value for money;
- Potential improvements (in particular if the policy has not performed as expected), and how the effectiveness of the policy could be maintained over time;
- An understanding of how transferable the evaluation results might be over context, place and time;
  The accountability of the delivering bodies, helping to satisfy requirements for external scrutiny and certain regulatory requirements.

Section 1: 'A visual guide to complexity'

Annex II presents these seven characteristics in more detail, plus an additional six that may also be of interest[14].

Note that in some cases there may well be overlap between the definitions, and/or a system could be described by multiple terms presented. The aim of this guide is to support readers in considering complexity within policy areas; and to provide suitable language for describing the complexity of the system they are working with.

---

[14] These additional six characteristics are: *domains of stability*, an *open system*, *distributed control*, *levers and hubs*, *nested systems*, and *multiple scales and levels*.

Section 1: Simple, complicated and complex evaluations?

In terms of policy-making and policy evaluation, complexity can arise at any one or more of multiple different levels. These include: the system in which the policy takes place (e.g. a complex socio-ecological system); and the policy itself (multiple actors, multiple actions).

For the evaluation, complexity can arise from any combination of the above sources and/or from the multiple and diverse stakeholder aims and perspectives involved. When designing evaluations, and seeking to determine whether complexity-appropriate approaches should be applied, it is useful to identify potential sources of complexity in both the system and the intervention:

- A **simple intervention** could have a single objective, a clearly defined target group and relatively few delivery activities, delivered by one agent, within a relatively short timescale. A **simple system** will likely be highly predictable, easily controlled, and with few actors affecting the system.
- A **complicated intervention** could be an intervention with multiple objectives, a programme of activities over different time periods, delivered by a number of different agents or require agents with specific expertise. A **complicated system** can have a large number of different actors, forces or relationships affecting how the system operates, but will be possible to predict with some confidence.
- A **complex intervention** and **complex system** may demonstrate complexity characteristics, such as non-linearity, adaptation and emergence. These characteristics are described in more detail below.

In practice, few policy evaluations may be categorised as 'simple', even where there is a single 'simple' intervention, because the environment in which the intervention operates usually forms part of a wider complex system.

Section 4: 'Complexity for Defra's policy domains - issues for designing'

There are a wealth of evaluation designs that work well with complexity, provided the evaluation is led and managed in a complexity-appropriate way.

Most of these designs are 'method-neutral' – meaning a wide range of different data-gathering and analytical methods can be used. These can include surveys, case studies, focus groups, randomized control trials and qualitative comparative analysis. There is no required use of particular methods.

There is no simple way to select the best design, and there is no gold-standard approach for complex evaluation. The choice will depend on the complexity of the intervention, characteristics of the system, evaluation purpose and the feasibility of the available designs and methods. However, tools to assist those who are designing evaluations in a complex context are discussed below. The design should seek to build on outputs and understanding of the intervention operation and context, developed at the 'Understanding' stage, such as system mapping.

Section 4: Monitoring, performance management and complexity evaluations

Effective evaluation requires accurate, timely and relevant data – this is particularly important for complexity-aware evaluations, where real-time data on project delivery activities, their impact, and any changes in the project's operating environment are vital inputs to the evaluation process.

To establish effective monitoring and evaluation frameworks, evaluators and project managers need to work together to identify:

- The key indicators required to monitoring inputs, activities, outputs, outcomes and impacts; and
- How, when and by whom data on change/s in key indicators are to be collected, verified, analysed, and reported.

This will help to ensure project monitoring, performance management, and evaluation reinforce each other.

Section 4: Enhanced discussion of methods

There are a number of tools, methods or approaches that can work well with complexity, such as:

- **Bayesian networks**, which use quantitative data and / or human perception to identify probabilities for key variables and so model complex systems. These models can be improved over time as new evidence emerges. Bayesian networks require technical skills (and specialist software can be helpful) to elicit and calculate probabilities, but can be used in policy appraisal and evaluation. For example, the document linked above shows how it has been used in tracking Bovine TB.
- **Agent based models** model the behaviour and interactions of individuals, households, businesses or other 'agents'. These models then enable evaluations to handle feedbacks and detailed interactions between agents. The approach cannot predict the future of a complex adaptive system, but can be used to offer insights into the range of possible futures, e.g. in relation to **adaptations in response to climate change**.
- **Qualitative Comparative Analysis** enables systematic comparison across cases (usually between 10 and 50). It takes account of 'complex causation', where a combination of 'attributes' may produce a given outcome. It is particularly useful when evaluators have a small to medium number of cases or interventions which are similar but applied in different contexts. The Environment Agency used this approach to navigate the complexities in waste crime policy.[15]

Each of these approaches, has its own strengths and weaknesses with respect to complexity, as do other methods or tools.[16]

---

[15] https://www.cecan.ac.uk/case-studies/environment-agency-enforcement-on-waste-crime

[16] For an overview of these strengths and weaknesses, see HM Treasury Magenta Book 2020 Supplementary Guide: Handling Complexity in Policy Evaluation

For further information on how to choose between methods, see The Magenta Book, plus **Befani, B. *Choosing appropriate evaluation methods*.** This resource provides a downloadable tool to identify appropriate evaluation methods. It uses a series of questions, to explore a method's ability to answer key evaluation questions (such as "What was the additional/net change caused by the intervention?"); and its ability to meet additional needs (for example the need to generalise evaluation findings outside the case/sample used for the analysis). It also explores the conditions that need to be met in order for a method to be applied in practice (for example "To what extent is it possible to control who does and doesn't receive the intervention?"). The tool provides a summary report on the most appropriate method/s given the information provided.

For further information on relevant evaluation and research methods and good practice, see:
CECAN Evaluation Policy and Practice Note Series (EPPNs)
CECAN syllabus: Qualitative Comparative Analysis; Systems Mapping; and Agent Based Modelling

Section 4: From design to commission

Key considerations when designing an evaluation, along with method and timescales, are proportionality and affordability.

A proportionate evaluation delivers findings that are of good quality and fit for purpose given the risks of getting the answer wrong.

Currently within Defra, there is no prescriptive guidance on the percentage of programme spend that should be dedicated to evaluation; decisions are made on a case-by-case basis and take account of issues such as the innovative nature of an intervention, the scale of an intervention, and the level of overall spend.

Once the complexity characteristics of the evaluation are identified, the range of potential approaches should be reviewed, considering the best options for a specific evaluation, based on feasibility and affordability.

Once a preferred evaluation approach has been identified, an estimate of the costs (including data collection, collation and analysis; project management; and reporting) should be produced.
If the available resources are not adequate for the preferred design, then adjustments to the design and/or to the resources available need to be made.

If additional resources cannot be secured, there are a number of ways to reduce costs, although each has potential implications for the evaluation outputs.

Ways to reduce evaluation costs while meeting evaluation objectives, and potential implications for the evaluation outputs:

- o Work with stakeholders to prioritise the evaluation questions to be answered (bearing in mind that losing any evaluation questions could have implications for the evaluation to meet its purpose);
- o Reducing sample sizes (this may result in reduced accuracy of estimates);
- o Reducing the number of 'waves' of research (again, this may lead to reduced accuracy)
- o Reducing the number of case studies to be undertaken (this may result in less depth of understanding of the system/intervention)
- o Embedding data collection in the overall management of the intervention (this may result in poorer quality data); and
- o Using alternative sources of data (this may have implications for data quality, accuracy or relevance).

See: Better Evaluation on implications of resources constraints: https://www.betterevaluation.org/en/node/5296

USEFUL QUESTIONS

- Is there flexibility to change the evaluation approach to respond to changing conditions? Have you considered the opportunities for flexible evaluation designs within the current parameters of commissioning rules? Explore the range of options through discussion with procurement colleagues.
- Have you identified a preferred approach based on the options available?
- Are the estimated costs of the preferred approach affordable?
- What adjustments to the design can be made to ensure the evaluation is affordable?
- What adjustments to the design can be made to ensure the evaluation is affordable?
- What are the implications of these adjustments (for example on data quality or accuracy), and have these been discussed with stakeholders?

CONSIDERATIONS FOR EVALUATION COMMISSIONERS

Complexity-aware evaluations may need to adopt flexible or iterative approaches. However, procurement rules may require external evaluation contractors to submit and deliver a fixed programme of work. This can present challenges for complexity-aware evaluations. The following suggestions may assist in maximising the opportunities for flexible evaluation design within common parameters of commissioning rules:

- Speak to procurement colleagues at the earliest opportunity when planning an evaluation to identify the degree of flexibility that can be built into the process and what options are available.
- Within tenders, recognise that circumstances may change. Tenderers could be asked to demonstrate how the evaluation approach can accommodate and adapt to changing circumstances. The evaluation purpose should be regularly reviewed.
- Consider building a feasibility or scoping period into the evaluation that can be used to define the following stages.
- Consider a staged approach to commissioning, with built-in decision points and options to reconsider the approach and/or contractor. Contractors can be asked to prepare the brief for the next stage of work as part of their remit. Where the evaluation is likely to be managed through a mix of internal and external resources, a break-point type approach could also be used; giving responsibility to a study

Steering Group as to whether to proceed at certain stages with external contractors or conduct the work internally.

- There may be a need to expand or revise the scope of future phases of an evaluation. This could also be achieved through commissioning additional studies, if helpful.
- Working with contractors on a 'co-creation' basis can sometimes be the best vehicle for delivering an evaluation that is sufficiently flexible and adaptive to cope with complexity. This can be built into contracts by stating that a co-creation work style is expected and/or by using a time and resources (rather than fixed price) contract style. Call off contracts can enable evaluation expertise and advice from policy experts and external consultants to be brought in at short notice.

Progressive thinking
No time for business as usual

We are a boutique economic development consultancy firm. We don't just offer strategies for economic development, we change the way our clients think about it. Bringing together a wealth of experience, we deliver a bold, rigorous, integrated vision of the future. We challenge assumptions, and leave our clients with the knowledge and capabilities to succeed.

Our offices

Manchester
61 Mosley Street
Manchester, M2 3HZ
+44 (0)161 261 9140

Leeds
67 Albion Street
Leeds, LS1 5AA
+44 (0)113 389 6400

London
28-32 Upper Ground
London, SE1 9PD
+44 (0)20 7910 5000

steer economic development

www.steer-ed.com