

The background is a vibrant mosaic of small, irregular tiles in various colors including red, orange, yellow, green, blue, and white. These tiles are arranged in a pattern of interlocking triangles. Overlaid on this mosaic are two large, semi-transparent circles: a light blue one on the left and a dark blue one on the right, which overlap each other and the text.

# The Complexity Evaluation Toolkit

In memory of Dione Hills.  
Dione was one of the founding members of  
CECAN and made immense contributions to our  
work, especially in grounding it in evaluation  
guidance and practice.  
She is very much missed by all those who  
worked with her.

## Authors

Martha Bicket, Brian Castellani, Corinna Elsenbroich, Nigel Gilbert, Dione Hills, Mark Gurney, Frances Rowe and Helen Wilkinson.

## Please cite this document as:

CECAN (2021) *The Complexity Evaluation Toolkit*. July 2021. Version 1.0.  
Online at [www.cecan.ac.uk](http://www.cecan.ac.uk)



This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

You are free to:

- Share - copy and redistribute the material in any medium or format
- Adapt - remix, transform, and build upon the material for any purpose, even commercially.

This license is acceptable for Free Cultural Works.

The licensor cannot revoke these freedoms as long as you follow these license terms:

- Attribution - You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- No additional restrictions - You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

## Contents

The chapter and sections headings below include anchors, so that if you click on them, they will take you directly to the appropriate Toolkit chapter or section.

<b>Chapter 1: Introduction to evaluation and complexity</b>	<b>04</b>
<b>Chapter 2: Commissioning a complex evaluation</b>	<b>07</b>
2.1 Assess the scope and scale of the evaluation	07
2.2 Specify the invitation to tender (ITT)	09
2.3 Develop and submit a response to the ITT	09
2.4 Evaluate bid	10
2.5 Initiate the evaluation	10
<b>Chapter 3: Designing an evaluation</b>	<b>11</b>
3.1 Developing a complex evaluation system	11
3.2 Establishing a Monitoring, Evaluation and Learning Framework (MEL)	11
3.3 Evaluability assessment	12
3.4 Selecting approaches for complex evaluation	14
3.5 Selecting methods	20
3.6 The challenges of data	21
<b>Chapter 4: Managing an evaluation</b>	<b>23</b>
4.1 Management of complex evaluations	23
4.2 Adaptive and agile management	23
4.3 Quality standards in complex policy evaluation	24
4.4 Data Quality	25
<b>Chapter 5: Achieving impact</b>	<b>27</b>
5.1 What happens after a complex evaluation, and why does it matter?	27
5.2 What are the barriers/enablers to complex evaluations being used and having impact?	27
5.3 Analysing past evaluations: an overlooked resource	30
5.4 Commitment to building capacity and capability in evaluation staff	30
<b>Appendix 1: Tools and resources</b>	<b>33</b>
<b>Appendix 2: Glossary of terms and acronyms</b>	<b>40</b>
Acronyms	42

# Chapter 1: Introduction to evaluation and complexity

*“How do you plan an evaluation when you have absolutely no idea what the outcome of an intervention will be since outcomes are emergent?”*

—Respondent to a CECAN survey of evaluators

**CECAN** is the result of the recognition that evaluation of policy interventions in a complex social world cannot be tackled by traditional evaluation approaches. CECAN is pioneering, testing and promoting innovative policy evaluation approaches and methods across *nexus* domains such as food, energy, water and the environment, through a series of ‘real-life’ case study projects. The Centre has developed approaches, practices and methods to improve evaluation in areas where issues of context, multiple perspectives, and emergence (i.e., the appearance of new and potentially unforeseen macroscopic properties in a system) are important.

This Toolkit collates the advances CECAN has brought about in the area of complex evaluation and covers key issues in commissioning, designing and managing an evaluation. The Toolkit is organised by chapters, each discussing the challenges complexity poses for different aspects of an evaluation and how these challenges can be mitigated, resolved or embraced. Whilst organised temporally according to the steps in an evaluation, the issues of complexity addressed in the chapters and the journey through them are intertwined and nonlinear. For example, whilst Chapter 2 is on commissioning—a somewhat logical place to start – scoping issues and understanding a project remain important across the evaluation process. Users may also need to revisit a chapter, as in the case of selecting methods, if a new system behaviour or evaluation needs emerge. Similarly, the issues in the impact chapter should not (and cannot) be left until the end. Because of this nonlinear overlap, there is an intentional degree of repetition in the Toolkit, as various issues resurface during the evaluation process. Functioning as a collation document, the Toolkit will also point to existing resources from CECAN and other organisations where more in depth knowledge is needed.

## **But what is complexity and how does it impact policy and policy evaluation?**

Complexity refers to the understanding that the social world is not mechanistic and entirely predictable but is made up of relationships and interactions between people and things, and whose actions are dynamic and overlap. Whilst there is not one agreed definition of complexity, a complex system is usually seen as a system consisting of

- diverse, interacting components
- interactions which lead to non-linear, non-proportional and emergent phenomena
- components that learn from and adapt to change



These features of complex systems lead to a lack of understanding of causal processes and a resulting unpredictability and uncertainty about the development of the system. The challenges to policy making and evaluation result from this uncertainty and unpredictability.

Although social systems are relatively stable over time, they are subject to turbulence and change and as such, this needs to be factored into policy evaluations. Policy making relies on understanding the outcomes of policy implementations. But if a system is complex and unpredictable, these outcomes may not be known. To mitigate this problem, policy-making uses evidence about past policy implementations and outcomes to inform future policy interventions. However, the nature of a complex system means that some kinds of knowledge about its past are not necessarily useful to understand its future. For example, if a policy intervention was implemented only in a small subset of the diverse components of a system (say a non-smoking intervention on a set of college students), the outcomes do not tell us how the policy might work on other components (the same intervention but applied to middle aged mothers). Even more problematic is that the components of a complex system learn and adapt to change. Responses to an initially effective policy intervention can change that strategy's ability to continue to achieve its goals, requiring a new set of interventions.

CECAN's goal is to help evaluators grapple with these challenges of complex systems, through developing methods that help to understand the causal mechanisms underlying changes in a particular complex system, to understand the inherent uncertainty under which complex systems need to be steered and to develop ways to improve future policy making by improving policy evaluation to give the right kind of knowledge to inform interventions.

**This Toolkit is aimed at those who have some basic knowledge and/or experience of evaluation and are seeking guidance and signposting on how to handle complexity.** These people may be policy analysts in government, evaluation consultants in the private sector, evaluators working in NGOs and charities and so forth.

The Toolkit draws from, summarises and builds upon concepts and guidance in key contributions that CECAN has made, such as the *Magenta Book Supplementary Guidance* (2020), and the *Complexity Evaluation Framework* (CEF) (2019).

This Toolkit signposts the reader to a range of evaluation resources. The full reference for each resource including their URL is detailed in **Appendix 1: Tools and resources** and each resource is listed in the order that they are mentioned in each Chapter. Some resources, bolded within the main text, include a hyperlink to their web location.

All acronyms and terms are fully explained in **Appendix 2: Glossary of terms and acronyms**.

**Chapter 2** looks at commissioning a complex evaluation. It walks the reader through the steps of commissioning, highlighting the importance of flexibility in responding to

emergent themes and the continuous communication between the commissioning organisation and potential evaluators.

**Chapter 3** covers a range of design aspects of complex evaluation, from a *Monitoring Evaluation and Learning framework* (MEL) and the importance of a Theory of Change (ToC), to discussing a range of approaches and methods including criteria for choosing between them.

**Chapter 4** considers the challenges of complexity to managing an evaluation. The focus is on adaptability and agility over the life span of the evaluation, discussing criteria and standards for quality assurance as well as the acknowledgement that some questions initially envisaged might not be answerable in the end.

**Chapter 5** focuses on the need and the difficulty of communication in and of complex evaluations. The chapter highlights the importance of continuous communication between different organisations, the need for timely communication of findings that might not be perfect but “good enough” to be informative. Finally, this chapter highlights the need for capacity building to build up cultures of complexity evaluation.

Running through all the chapters are some key messages to keep in mind when considering a complex evaluation.

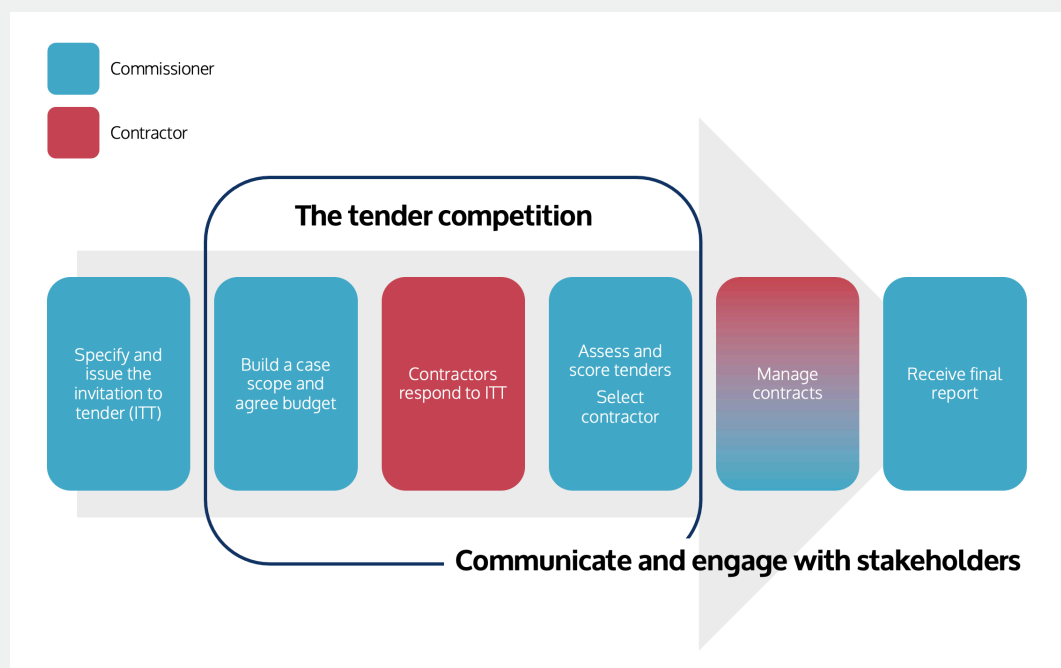
- **Start early.** Well-designed evaluations sit along-side policy implementation and inform ongoing design. Appraisal and evaluation become a continual process.
- **Use a Theory of Change (ToC).** A ToC will help to share understanding between stakeholders and inform planning and capture emerging understandings.
- **Don't aim for perfection.** Aim for ‘good enough’ - there is no such thing as perfect in a complex and changing environment. Timeliness of findings that are ‘good enough’ to inform policy going forward generates value - waiting for a perfect answer that never comes or is too late does not.
- **Stay flexible.** Select the approach that best fits your requirements from the evaluation and the complexity inherent in the system. Be prepared to change and adapt your ToC and plans to what you find, and to changes in the policy and external environment. Be prepared to adopt different ways of doing things.
- **Communicate widely and often.** Work with your stakeholders, get them engaged in the evaluation from the start, keep them engaged throughout - including those who are best placed to use the findings.
- **Plan for change.** Commission for all the above. Don't use the same old procurement templates if they won't deliver these essentials.

## Chapter 2: Commissioning a complex evaluation

The first step in organising an evaluation is to *commission* it. Often, the intention is to commission an outside and independent evaluation practitioner or consultancy to do the work, but sometimes evaluations are done ‘in-house’ by the organisation’s own staff. In either case, it is necessary to be clear about what needs to be done and to what time scale, and this is best achieved by developing a *specification* of the evaluation, which may form the basis of an invitation to tender (ITT) or request for proposals (RFP) for potential evaluation teams to consider.

In this section, we outline the steps involved in commissioning an evaluation and the decisions that need to be made along the way, with an emphasis on the issues that need to be considered when commissioning complex evaluations. For further information, see the tools and resources for Chapter 2 in **Appendix 1: Tools and resources**.

Figure 1 is a visualisation of the commissioning process for an external contractor.



**Figure 1:** Simplified description of the different stages in the lifecycle of an externally commissioned evaluation. Derived from Cox (2019:3).

### 2.1 Assess the scope and scale of the evaluation

The first step in commissioning an evaluation is to define what is to be evaluated: the **scope of the evaluation**. This can be tricky in complex policy domains. There is a temptation to consider that everything depends on or is influenced by everything else, and so the scope must be all-encompassing. Careful consideration of the objectives of

the evaluation and the needs of those who will use its results may help to narrow the scope to something feasible and reduce the risk of creating a large number of evaluation questions, some of which may not be answerable in a complex environment.

This leads to the question of the **scale of the evaluation**: is it to be detailed and wide-ranging, involving much primary data collection, or 'light-touch', or something in between? One answer will come from the complexity of the policy being evaluated and the size of the budget allocated to evaluation. Some large-scale programmes, may have sub-parts and decisions will need to be made about whether the overall programme or some/all of the sub-parts will be the subject of the evaluation. This will enable a sense of the data (qualitative and quantitative) that will need to be obtained in the course of the evaluation.

Another consideration is the **timing**: an evaluation needs to be completed in time for its findings to be acted upon. If, for example, a policy change is made before the results of an evaluation of the policy are reported, the evaluation will have very limited worth. This implies that evaluation in fast-moving policy areas needs to be quick or designed to be very flexible and able to track policy change.

**Decisions on scope and scale** need to be made in consultation with the budget holder, policy development lead and the policy implementation manager. Convening a stakeholder group, including all those with an interest in the evaluation or their representatives, at a very early stage is vital. It is important that the evaluation is included as part of the overall project plan and capacity is available to implement the findings.

Preliminary work may be needed to work with **stakeholders** to create space to think about what the desired policy outcomes are and to counter linear ways of thinking. This is where the anticipated outcomes of the programme as a **Theory of Change** (ToC) can be discussed, modified or developed.

The role of the **stakeholder group** can include assisting in development of the specification, receiving reports on progress, and eventually taking account of the evaluation results in their own work. It can also be responsible for feeding back into the evaluation as it is ongoing by commenting on data sources and participants, emerging findings, and providing a steer on the direction of the evaluation.

Commissioning an evaluation can take longer than you might expect. Key factors to take into account include

- Time to write the specification
- Consultations with budget holders and the stakeholder group.
- Tendering timescales: submission of proposals, screening and interviewing
- Ethical clearance (if required)

To summarise, the initial steps to be taken before one can start commissioning include

1. Clarify what will be evaluated
2. Describe the Theory of Change (ToC)
3. Identify who are the primary intended users of the evaluation and what will they use it for



4. Develop agreed key evaluation questions grounded in the purpose of the evaluation and an understanding of how findings are to be used
5. Decide the timing of the evaluation
6. Identify what resources are available for the evaluation and what will be needed
7. Identify the staff time and commitment required to participate in the evaluation, e.g., in interviews, gatekeeping roles etc
8. Clarify any data access issues that will need to be addressed in order for the evaluation to begin

## 2.2 Specify the invitation to tender (ITT)

Most organisations will have standard invitation to tender processes in place, which are used to commission evaluations or for the purchase of products and services. However, because of the challenges of tackling complexity in evaluation, it may be necessary to adapt these processes to cater for flexibility, responsiveness, the ability to engage different actors effectively and to communicate well at all stages of the evaluation.

The most important elements of an ITT are the project specification and the selection criteria to be used to choose which bid to accept.

The project specification should describe

- The context
- The purpose of the evaluation
- Any constraints (e.g., the availability of staff for interviews; the timing of data collection, interim and final reports)
- The intended use of evaluation outcomes

It is often a good idea to hold a meeting open to all prospective bidders, timed well before the deadline, at which the context of the evaluation can be explained and requests for clarifications addressed. This is important as marking the start of a dialogue with the potential evaluators, a dialogue that should be continued throughout the project.

## 2.3 Develop and submit a response to the ITT

From the perspective of bidders, bidding against an ITT for a complex evaluation can be a challenge. The bid should aim to be convincing, properly costed and clear about what is to be done and when. On the other hand, there may be much uncertainty about how the evaluation should proceed, what factors are important to measure, and what skills will be required, all making the bid hard to define.

Typical bids include

- An introduction, which outlines the bidder's understanding of the project and what is being evaluated and why
- A section on the methods to be deployed, and how those methods will enable the evaluation to be carried out effectively
- A demonstration of the bidder's track record, with examples of similar projects

that have been carried out successfully, and the lessons that have been learned from those that can now be used in the proposed project

- Testimonials or references provided by organisations that the bidder has previously worked with

## 2.4 Evaluate bids

Once the bids have been received, they need to be assessed against the criteria established when the ITT was specified. This can be a difficult task when the evaluation is complex: bidders may have proposed very different approaches.

The assessment panel needs to bear in mind

- The intended outcomes of the evaluation and the extent to which the proposed methods will achieve these
- The track record of the bidders - have they been successful in previous projects?
- The extent to which they have understood the brief and especially the need to take account of complexity
- The extent to which they are prepared and able to work flexibly, responding to emerging themes and issues

## 2.5 Initiate the evaluation

Typically, an inception meeting marks the initiation of the evaluation and would include

1. A reiteration of the aims and objectives of the evaluation
2. Reviewing the proposal submitted, and discussing any potential amendments
3. Reviewing timescales for the evaluation and ensuring they are realistic
4. Providing information needed by the evaluators such as key contacts and essential data
5. Agreeing the communication protocols and management arrangements
6. Agreeing a schedule of deliverables, such as progress and final reports
7. Agreeing the contract and an invoicing schedule
8. Making the evaluator aware of any sensitivities, concerns and risks that could have an impact on the successful delivery of the evaluation
9. Confirmation of data access
10. Ensuring that the evaluation timescales are congruent with the overall project/ implementation plan

Commissioners should identify a named person as the key contact for regular communication, setting up day-to-day arrangements and so forth. In addition, there may be a more senior person who maintains a strategic overview of the evaluation, chairs the stakeholder group and formally accepts the deliverables.

## Chapter 3: Designing an evaluation

### 3.1 Developing a complex evaluation system

Evaluations do not take place in a vacuum: they are part of a wider system of planning, developing, monitoring and assessing an intervention or policy, as well as generating learning and evidence for future interventions. In order for an evaluation to embrace complexity the larger system in which it takes place must also embrace and comprehend complexity.

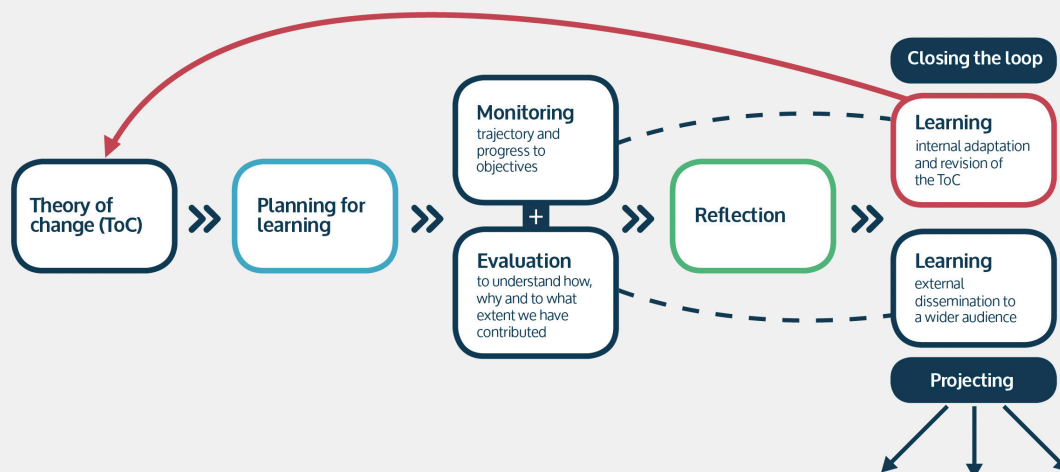
In thinking about this, it is helpful to consider what a “complexity-appropriate” *Monitoring, Evaluation and Learning (MEL) framework* might look like and how one goes about not only engaging in such an evaluation but also engaging the wider system in which an evaluation is taking place to think in similar complex systems terms.

For further information about the issues and resources discussed in this chapter, see the corresponding tools and resources for Chapter 3 in **Appendix 1: Tools and resources**.

### 3.2 Establishing a Monitoring, Evaluation and Learning Framework (MEL)

A *Monitoring, Evaluation and Learning (MEL) framework* provides a way of bringing both the management of policy intervention activities and the evaluation into one overarching strategic plan, which includes plans for how data and findings from these feed into learning, so contributing to the overall outcome.

A helpful resource for developing a MEL framework is the UK Government Stabilisation Unit’s guidance note on ***Monitoring, Evaluation and Learning (MEL) in Conflict and Stabilisation Settings (2019)***.



**Figure 2:** Schematic MEL Framework. **Source:** Stabilisation Unit (2019, slide 37)

The guidance note sets out a number of elements required for a MEL framework, with detailed information about each element (in an interactive format). A schematic outline of this process is shown in Figure 2 (pg. 11).

The Stabilisation Unit's guidance suggests that an important starting point is the establishment of a Theory of Change (as we noted earlier in Chapter 1), from which a 'learning plan' is developed identifying how monitoring and evaluation activities will feed into reflection and learning, both for adapting internal plans (and revising the ToC map) and sharing with a wider audience.

Other useful points include

- Clarifying the difference between evaluation, research and monitoring
- Being clear about the values and principles underpinning the MEL framework
- Identifying risks and risk mitigation strategies
- Defining realistic results (i.e., what can be directly controlled or influenced)
- Identifying a 'results framework' and the data required to test this
- Selecting appropriate evaluation designs
- What is needed to support learning

When establishing a MEL, it is important to ensure that both the monitoring activities and the evaluation activities are complexity-appropriate. In some cases, for example, extensive Key Performance Indicators (KPIs) may have been established for an intervention to be used as the basis of monitoring, before the full complexity of the intervention and its setting is understood. Some of these might later turn out to be totally inappropriate.

The Stabilisation Unit's guidance, while useful, does not discuss in-depth how to establish and manage the MEL. For these steps, we recommend **BetterEvaluation's guidance** on how to *Manage an evaluation or evaluation system* (BetterEvaluation, 2013), which covers key points related to the design and management of an evaluation system, including

- Consulting stakeholders (about evaluation aims, interests and roles)
- Establishing decision making processes
- Determining and securing resources
- Document processes and agreements (e.g., contractual arrangements)
- Developing evaluation capacity

### 3.3 Evaluability assessment

For many years, evaluability assessments have been used, particularly in the international development field, to identify 'the extent to which an activity or project can be evaluated in a reliable and credible fashion' (OECD-DAC 2010; p.21).

Undertaken as part of designing an *evaluation strategy*, evaluability assessments involve a review of documentation and conversations with key stakeholders. They

help to clarify not only the feasibility but also the purpose and timing of the evaluation. They assess

- Evaluability “in principle”, given the nature of the project’s Theory of Change (ToC)
- Evaluability “in practice”, given the availability of relevant data and the ability of management systems to provide it
- Evaluability “in value”, relative to the utility and practicality of an evaluation, given the views and availability of relevant stakeholders

There are a number of resources that help with undertaking an evaluability assessment including checklists and guidance on their use. Including the *Evaluability Assessment Guidance* (BetterEvaluation, 2014).

### 3.3.1 Evaluability assessment in complex settings

In a complex intervention or setting, an evaluability assessment helps key stakeholders reach a common understanding about the complexity of an intervention and the system it is working within, and how this can be addressed in the evaluation design. Reviewing the strength and limitations of different evaluation designs can help inject a level of realism both about the potential costs (and expertise) required for the evaluation, and the kind of findings the evaluation will generate.

While an evaluability assessment can be undertaken at various points in the planning of an intervention, it is generally recommended to undertake this early, as this helps to design the intervention in a way that aids its evaluation. This is compatible with the broad message in this Toolkit, that evaluation design should begin early in the planning of a policy intervention.

- Evaluability assessments involve consulting stakeholders about the purpose and feasibility of an evaluation. This fits with the general message about taking into account different perspectives in a complex system and helps ensure that complexities ‘on the ground’ are recognised and taken into account at an early stage.
- Evaluability assessments require a Theory of Change (ToC) to be drawn up (see section 3.3.2) and can help clarify the boundary around the intervention. Establishing boundaries in a complex setting are challenging because complex systems are ‘open’ to influences from their wider context. Drawing up a ‘system map’ prior to drawing up the ToC map can be helpful in these circumstances, ensuring that key features of the wider system are recognised and taken into account.

A key resource for evaluability assessments and complexity is the ***Magenta Book Supplementary Guidance*** (2020). This supplement, authored by CECAN, aims to “explain what complexity thinking is, what the features of complex systems are, and how new methodologies and tools can equip policymakers to work with unavoidable complexity” (p.8). See **Appendix 1: Tools and resources** for more information.



### 3.3.2 Theory of Change for complex settings

A key concept throughout this document is “Theory of Change” (ToC). ToC is a common tool used in evaluations. In fact, a ToC serves as a crucial first step in any complex evaluation in its own right, rather than as part of an evaluability assessment.

The ultimate intention of the ToC is to describe how change is expected to happen; it describes the pathways to change and how activities are understood to produce a series of results that contribute to achieving the final intended impacts.

In the context of complexity, combining system mapping approaches with a ToC enables stakeholders, evaluators and commissioners of evaluations to develop frameworks that take into account key features of the whole system that are likely to influence the outcomes of interventions. These features include, for example, feedbacks and interactions with other policies and the institutional context.

For a useful discussion of the use of ToC in evaluation, an approach for building system-based (complexity-appropriate) ToC using participatory system mapping (PSM) and a practical guide to PSM see **Appendix 1: Tools and resources**.

## 3.4 Selecting approaches for complex evaluation

The *Magenta Book Supplementary Guidance (2020)* gives detailed guidance on how to choose a complexity-appropriate approach (or combination of approaches) for a particular evaluation. This section presents key aspects of that guidance. For further information, including helpful tables for use in planning complexity-appropriate evaluation approaches, see chapter 4 of the *Supplementary Guidance (2020)*.

An important step in the design of the evaluation is selecting the most appropriate approach given the aims of the evaluation, the data available or that can be collected and the resources available. In complex environments this may not be a trivial task. Here we provide guidance for evaluation commissioners and practitioners on how to go about selecting the overall approach and signpost useful tools for selecting specific data gathering and analysis methods to deliver the approach.

Relative to our guidance, it is important to remain flexible about how one goes about the process, as requirements from the evaluation may change over time and detailed methodological requirements may emerge as understanding improves. (Further guidance on how to adjust to and accommodate such changes over time in the commissioning processes and in management of the evaluation is presented in Chapter 2 of this Toolkit.)

### Key points

- There is a wealth of evaluation approaches and methods available that work well with complexity, particularly when implemented within a Theory of Change (ToC) framework.
- The approach chosen will depend on the complexity characteristics of the system, the evaluation purpose and the feasibility of the available approaches.

- Particularly useful ways of framing evaluations for complex settings include
  - **Participative approaches**, including system mapping, which can bring stakeholders together to generate deeper, shared understanding and provide a safe space in which participants can expose and air differences
  - **Developmental approaches** that involve stakeholders in the evaluation and as agents for change and are particularly useful for supporting adaptive management approaches (structured, iterative decision-making processes in uncertain environments)
  - **Qualitative, theory-based approaches**, which can be used to explore whether the policy is contributing to change, in what way, and the underpinning mechanisms of change, to provide rich information and potentially useful lessons for similar policies and contexts
  - **Configurational (case-based) approaches**, which help identify those factors, or combinations of factors, that appear necessary or sufficient to success including contextual factors
  - **Computational system modelling**, which can provide a ‘virtual’ *counterfactual*, a vision of what might have happened in the absence of the policy when it is not possible to establish an experimental *counterfactual* and allow the evaluator to project forward into the future and explore what further change may happen
- Detailed methodological requirements may only emerge over time. Flexibility is therefore required, with evaluators and commissioners regularly reviewing the design to determine how well it is working and whether it should be modified.

### 3.4.1 The different types of approach

There are a number of ways we can approach an evaluation and a number of different ways these overall conceptual approaches can be categorised.

Here we have chosen to cluster approaches into six groups, differing mainly in how they establish conclusions. These are

- **Participatory, emancipatory and adaptive approaches** such as developmental evaluation, action research and peer challenge, highly responsive and exploratory approaches in which stakeholders take an active part in the delivery of the evaluation providing real time feedback on the policy. Participatory methods are particularly recommended for constructing the Theory of Change (ToC) especially where there are multiple, perhaps conflicting, viewpoints and where there are high degrees of uncertainty.
- **Theory-based approaches** which articulate a theory of how the policy is working to deliver change, then seek to test this to investigate whether, why or how the policy causes or contributes to observed results, and whether alternative explanations can be ruled out. Approaches include
  - *Systems mapping and modelling*, used to generate, progress and test the Theory of Change (ToC) through an iterative process of developing and testing a formal model of the system

- *Generative causation* approaches, which seek to articulate the underlying mechanisms or processes of change. They test the theory empirically to investigate whether, why or how the policy causes or contributed to observed results, and how context influences these. Examples of *generative causation* approaches are realist evaluation and contribution analysis
- **Configurational case-based approaches** such as *qualitative comparative analysis* (QCA), which support systematic analysis of a number of interventions being evaluated. They identify the configuration of factors, or combinations of factors, that appear necessary or sufficient, to success
- **Counterfactual approaches**, including:
  - Experimental approaches such as randomised control trials, and quasi-experimental such as difference in difference (DiD), which provide a usually quantitative measure of the extent to which any observed changes in an outcome of interest were caused by the intervention (or treatment) by means of a comparison of results obtained for a treatment group with those in a non-treatment control
  - Predictive approaches, which attempt to predict what would have happened in the absence of the treatment using statistical or simulation modelling; in its simplest form, predictive approaches use informant opinion to understand whether impacts would have happened in the absence of an intervention
- **Statistical association approaches**, that look for correlations between cause and effect or between variables, to explore the influence of (usually) isolatable multiple causes on a single effect, while controlling for *confounding factors* that might suggest a spurious association
- **Synthesis designs**, such as realist synthesis, which seek to draw conclusions by combining results from evaluations drawn from several contexts

For all of these approaches and for complex settings we recommend that they should be applied within a Theory of Change (ToC) framework as discussed above in section 3.3.2.

They can all be more or less participatory in nature, that is stakeholders can be more or less closely involved in the design and delivery of the evaluation. In complex settings involving many stakeholders with differing perspectives, participatory methods are particularly useful and may be essential.

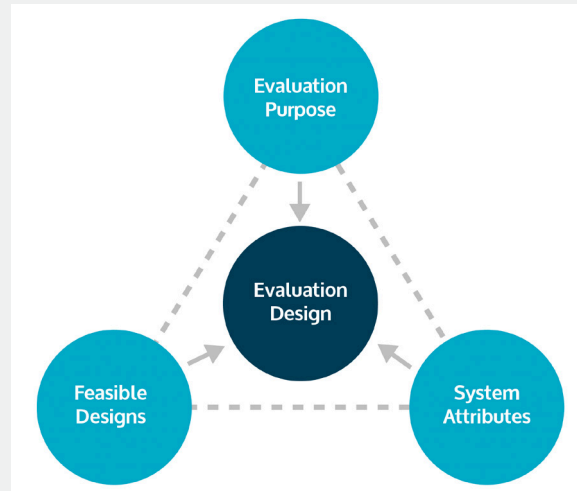
None of these approaches is mutually exclusive. In evaluation it is common to combine different methods so that conclusions are based on several different sources of information gathered in different ways. For complex evaluations, often what is required is a hybrid design in which two or more overall approaches are combined and tailored to meet the needs of the evaluation.

### 3.4.2 Selecting the most appropriate approach - the design triangle

Many of the familiar evaluation methods and approaches described in the **Magenta Book** (2020a) can generate useful findings as long as the complexity of the policy, and its setting, is acknowledged, and the findings treated appropriately.

There is no simple, mechanistic way to make design decisions. The design triangle below illustrates three interrelated key considerations to take into account when selecting an evaluation approach: the purpose of the evaluation, the attributes of the system being evaluated, and the feasibility of possible designs (Figure 3).

The 'design triangle' illustrates three inter-related factors that should be considered in establishing or reviewing an evaluation design. The diagram emphasises that many of these decisions are interconnected. For example, the kinds of evaluation questions that can be asked partly determine the selection of approach and methods, which also has to take account of system attributes in understanding the kinds of questions that can be answered.



**Figure 3:** The Design Triangle. Source: CECAN (2020)

The next section provides guidance on how to apply these three key aspects of design.

### 3.4.3 Evaluation purpose

Evaluations are undertaken for a range of purposes, including accountability and learning, listening and building: ensuring diverse voices are heard and building trust and legitimacy across stakeholders. The purpose will be reflected in the evaluation questions that key stakeholders want to see addressed. In a complex policy situation, there may be a gap between what key stakeholders would like to know, and what it is possible to conclude from the evaluation.

Where the policy is cross-sectoral and/or involves several different stakeholder groups, differences can also emerge between these about the purpose of the evaluation, and how important it is to have

- A learning element to the evaluation (as well as demonstrating accountability and evidence of impact) – Learning aims are particularly important in the evaluation of a complex intervention
- A process element in the evaluation, as well as assessing outcomes and impacts. **The Magenta Book** (HM Treasury, 2020a) distinguishes between process, impact and value for money evaluation. When a policy or its setting are complex, a process element to the evaluation will be important for tracking changes to the implementation process over time, helping to explain why impacts have, or have not been achieved

Commissioners should be prepared to explore not just how well the intervention is working and how it can be improved, but also question whether quite different approaches may have produced better results. In this way, appraisal and evaluation will merge into a continuous process of learning and policy evolution. For more, see Chapter 4 in the *Magenta Book Supplementary Guidance* (2020).

### 3.4.4 System attributes

As has already been noted, a key stage in planning an evaluation is to understand the intervention itself, and the setting into which it is being delivered, ideally through some kind of system, logic or theory mapping. Knowing about the characteristics of the system will be an important element in choosing an evaluation approach.

Not all of the characteristics will be apparent at the outset of an evaluation. Different elements of complexity (*feedback loops, sensitivity to context, levers and hubs*) will often emerge over the course of policy delivery, or as the evaluators find out more about the setting in which it is being delivered. This may be unwelcome news if those who planned the policy were assuming a generally 'linear' model of policy implementation and can also throw an inappropriate *evaluation strategy* off course.

The involvement of stakeholders at the start of the evaluation planning process who have experience or 'local' knowledge about the kind of intervention being used and its setting will be particularly helpful. In policy arenas where understanding about complexity is still at an early stage, it can be useful to engage experts with an understanding of complexity to advise on planning the evaluation.

Table 5 in the *Magenta Book Supplementary Guidance* (2020) highlights some of the key complexity challenges and approaches that may be selected to mitigate them.

### 3.4.5 Feasible designs

As well as considering which approaches are best suited to the aims of the evaluation and the level of complexity in the system, the resources available (and proportionate) in terms of funding, time, data and skills must also be taken into account.

#### **Acceptability**

An important consideration can be how acceptable the approach is to key stakeholders and the resources required to facilitate the evaluation. There is now a growing body of approaches becoming available that are particularly useful in addressing complexity, although knowledge of these remains patchy. Adoption of appropriate methods can be particularly challenging if the prevailing culture favours evaluation approaches that do not address the challenges of complexity very well.

The following issues may be particularly challenging for some stakeholders

- **Evaluator objectivity:** some of the methods described are more explicit in their acknowledgement that evaluators become part of the situation as they interact with people on the ground



- **Participatory evaluation:** approaches can be particularly helpful in enabling local stakeholders to play a role in the design and 'sense making' aspects of the evaluation but may be unfamiliar to some stakeholders
- **The relative value of qualitative and quantitative data:** several of the approaches rely on qualitative as well as quantitative data; this will be challenging in policy environments, or with stakeholders, that view only quantitative findings as robust
- **The level of precision and detailed explanation that can be achieved:** the level of quantitative rigour and certainty of outcome may be limited, even when using sophisticated evaluation methods

Strong leadership from project managers, a clear explanation about why a particular approach is being adopted, and the opportunity to air differences of perspective will be important.

As the approaches and methods used may need to change as understanding of the system and intervention develops, and as new players become involved over time, this process of communication needs to continue throughout the evaluation.

### **Resources**

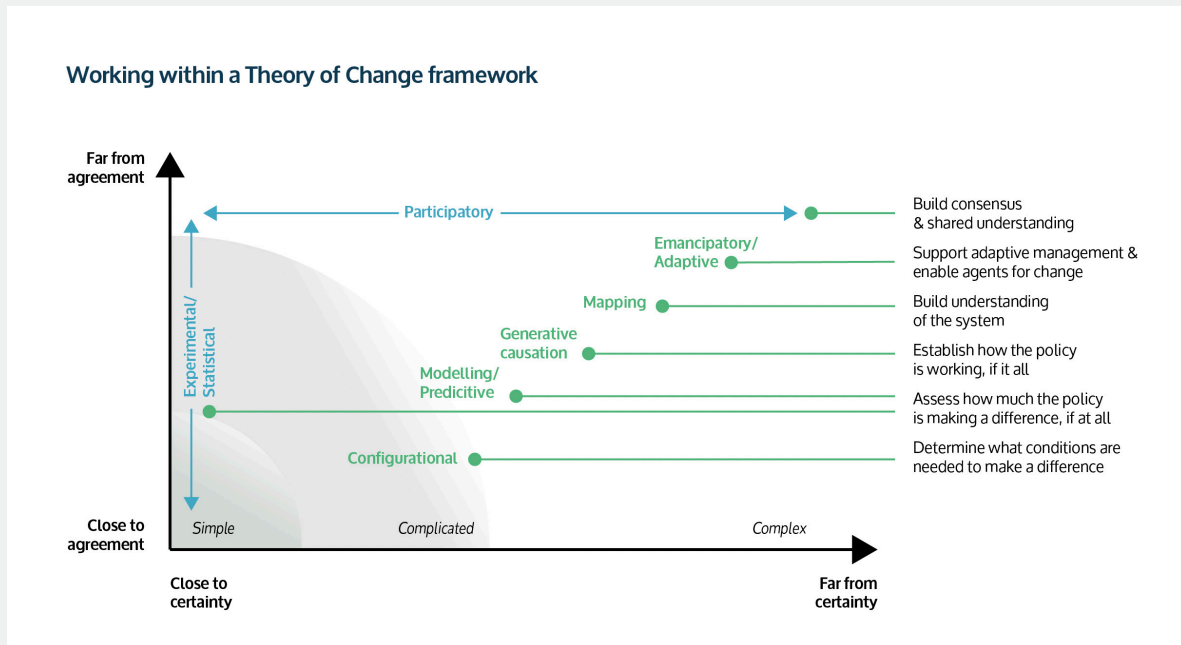
Resources will be an important consideration in deciding how feasible an approach is. Even in a relatively straightforward evaluation design, issues related to expertise, access to data, timing and the amount of time needed may require additional thought, as well as the overall budget. For more about data see the section 3.6 on data below.

Table 6 in the [Magenta Book Supplementary Guidance \(2020\)](#) summarises the factors relating to each type of approach affecting feasibility, including the specialist skills and levels of resource required.

### **3.4.6 Bringing it all together**

The diagram in Figure 4 (pg 20) shows how consideration of the purpose of the evaluation and system attributes relevant to complexity inform which of the available approaches are likely to be most useful, and feasible. So, for example

- Where the policy is relatively simple (left hand portion of the figure), randomised control trials, quasi-experimental or statistical approaches may provide information on whether, and to what extent, the policy worked. To understand why the policy worked and what aspects of it may work elsewhere, it will be necessary to consider using e.g., *generative causation* or configurational approaches as well.
- Moving towards the upper right-hand area of the figure, where there are low levels of understanding and agreement about the system and the policy, participatory system mapping will be useful to help develop a common understanding of the system and how the policy is intended to deliver impact (the Theory of Change), adaptive approaches can be used to generate information to support Adaptive management in a timely manner.



**Figure 4:** Identifying potentially useful approaches. Source: CECAN (2020)

## 3.5 Selecting methods

Many of the approaches described are to a large extent 'method neutral'. That is, a range of different methods can be used to gather and analyse evaluation data, within the overall framework provided by the selected approach or approaches.

There has been a lot written on the various methods available for engaging in computational social science, complexity methods, and complexity in evaluation. It can be overwhelming, particularly given the level of expertise necessary to use some of these tools and the time needed to invest in learning them. As such, it is important to have a way to make sense of these various methods to save time and effort and pick the right method for the right job.

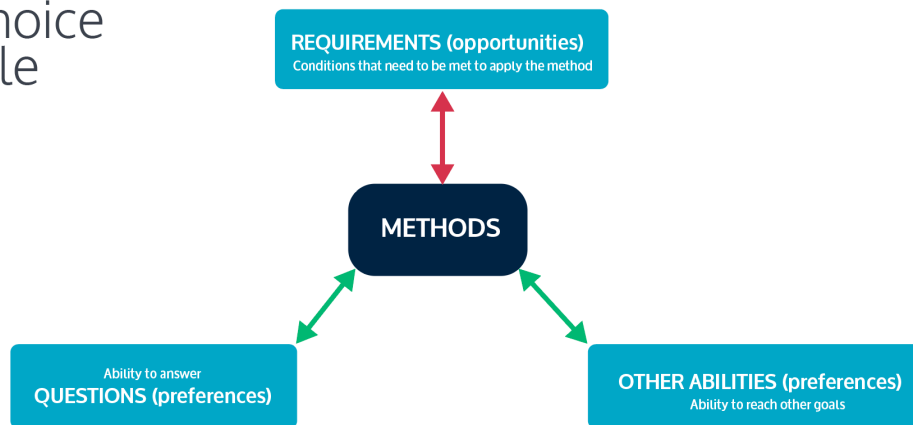
### 3.5.1 Choosing Appropriate Evaluation Methods – a tool for assessment and selection

The ***Choosing Appropriate Evaluation Methods tool*** was originally developed by Barbara Befani and Michael O'Donnell.

The Excel-based tool and accompanying report aim to support the process of methodological selection by helping users to make an informed and reasoned choice of one or more methods for a specific evaluation. The aim is not to necessarily provide a simple answer, but to refine, clarify and articulate the reasoning behind choice and have both commissioners and evaluators weigh pros and cons of possible options in a logical and structured way.

The tool complements the guidance in this Toolkit on methods and uses a revised design triangle, called the choice triangle, to help the user narrow down and identify suitable methods (see Figure 5).

## The Choice Triangle



**Figure 5:** The Choice Triangle. Source: Befani (2020)

The tool takes the user through a series of questions that help identify candidate methods based on each of the three elements of the choice question

- The method's ability to answer key evaluation questions
- The method's ability to carry out specific tasks/ achieve specific goals
- The team's ability to accommodate the method's requirements

The tool then indicates which methods best match the users need. For an example of the questions used in the tool, see the report: ***Choosing Appropriate Evaluation Methods – A Tool for Assessment and Selection*** (Befani 2020).

### 3.6 The challenges of data

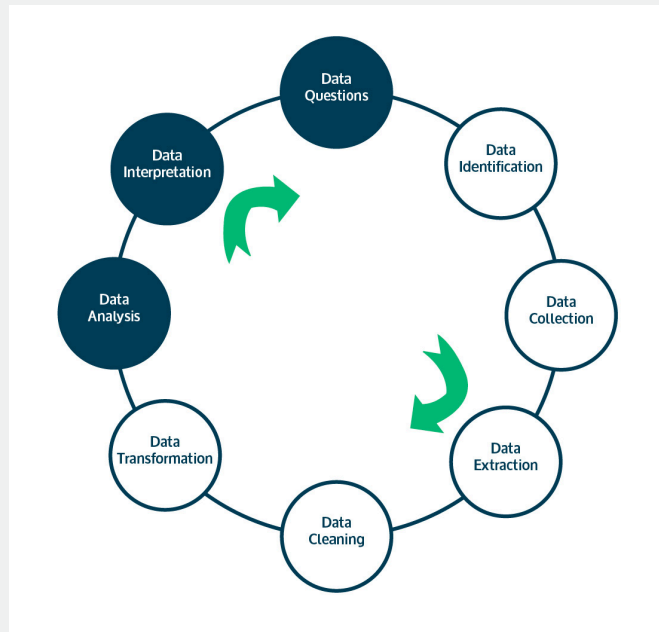
The approaches described above use qualitative or quantitative data or both and each presents specific challenges. These challenges can be organised into two major categories: what data do you need and how do you extract information useful for complexity-based evaluation from existing data?

- Picking the right data – understanding your clients' opinions and perspectives is one of the most important parts of the evaluation process: this may involve using technology or paper surveys for data collection needs, along with *Participatory Systems Mapping*, focus groups, interviews etc.
- Much data today is not natively in structured format; for example, tweets and blogs are weakly structured pieces of text, while images and video are structured for storage and display, but not for semantic content and search. Therefore, transforming such content into a structured format for later analysis is a major challenge
- Many evaluation studies data are not sufficiently longitudinal in nature, with at most pre- and post-test/intervention data. And very little of it is in real-time

- The other major issue is making use of “big data”, which is widely available, but also presents its own challenges

### 3.6.1 Extracting knowledge

One approach to data management, particularly for complex evaluations, is data mining (Figure 6).



**Figure 6:** Data Mining Diagram. Source: Castellani and Rajaram (2021)

Data mining typically involves an ongoing process of data collection, management and analysis that seeks to uncover non-obvious patterns in large databases for the purposes of knowledge discovery. The emphasis here is on the non-obvious.

Concerns about theoretical data dredging aside, a lot of data mining is resolutely exploratory in nature, seeking to know what we don't know, as well as figuring out the unknowns. It is also about conducting 'explorations' of large amounts of information.

## Chapter 4: Managing an evaluation

### 4.1 Management of complex evaluations

The task of managing an evaluation is inherently complex because it involves working between multiple stakeholders, often involving multiple academic disciplines, professional groups, sectors and government departments. In a complex intervention, the complexity of the task is increased because the intervention involves multiple levels, individuals and organisations.

Except in the case of self-evaluation, there are two managers in an evaluation: the commissioning manager and the manager of the external evaluation task and team.

The commissioning (or project) manager works at the interface between the end users of the evaluation findings (in government departments, the policy maker and sometimes policy makers in other government departments), and other key stakeholders (including those commissioning and delivering the intervention) and the contracted evaluator.

The evaluation manager works at the interface between the commissioning manager, those funding and implementing the intervention on the ground, their evaluation team and their own organisation.

Many of the points made in Chapter 2 relating to commissioning are also relevant to the management of complex evaluations. There are also some additional points in chapter 3 of the *Magenta Book Supplementary Guidance (2020)*. Particularly important points include:

- Having an adaptive or agile approach to management
- Applying appropriate quality standards
- Ensuring that there is capacity, and capability, in the system to commission, undertake and use findings from a complex evaluation

Section 3.6 of the Supplementary Guidance provides a series of useful questions for commissioners and managers in the evaluation planning and delivery process.

For further information about the resources discussed in this chapter, see the tools and resources for Chapter 4 in **Appendix 1: Tools and resources**.

### 4.2 Adaptive and agile management

Traditional management approaches emphasise adherence to detailed plans to solve problems in a linear and mechanistic way.

In complex policy environments, with multiple stakeholders and high levels of uncertainty a more agile and iterative process is required. This requires greater flexibility in the management and organisational response to the evaluation than may be familiar.



It typically involves

- A more open specification for the evaluation
- Consensus building with a range of stakeholders
- Flexible governance and management of the evaluation in response to emergent change, responses to the intervention and the impact of the evaluation process
- Regular review points throughout the evaluation

The *Complexity Evaluation Framework* (CEF) (2019) is a framework of key considerations to guide the scoping, commissioning, management and delivery of complexity-appropriate evaluations. It champions an adaptive approach to complex evaluation which is iterative and embedded throughout the policy cycle. This approach acknowledges that aspects of an evaluation may be subject to some degree of change over the course of the evaluation, for example as stakeholder understanding increases or as their objectives change.

The CEF highlights that two key components of an evaluation, (i) **scoping/** understanding the system and intervention and (ii) **designing the evaluation**, may both continue to develop and need to be updated throughout the evaluation (for example, as an intervention is evaluated, more will be understood about the intervention and any new changes in its context, and therefore how best it can be evaluated). These activities are conducted with the ongoing engagement of stakeholders and understanding, and learning are fed back and embedded into relevant processes both inside and outside of the evaluation, and not just at the end.

For more on adaptive management as an approach to working with uncertainty and complexity, see the ***Bond Guide on Adaptive Management (O'Donnell, 2016)***.

### 4.3 Quality standards in complex policy evaluation

Commissioners of complex evaluations may struggle when assessing the quality of evaluations since there are, as yet no established and generic standards for ensuring that complex evaluations are of an appropriate quality.

Most quality standards are drawn from research standards and experimental designs and are not easily applicable to complex settings. For a review of some of the recently developed standards of evidence see *Mapping the Standards of Evidence used in UK social policy* (Puttick, 2018).

Although there are no quality guidelines specifically for complex evaluations, there are a number of generic quality standards that are helpful.

The key principles for a good quality evaluation include

- Has taken steps to understand the intervention – and recognised what level of complexity is present early on, and designed an evaluation that is appropriate for this level of complexity
- Has involved consultation – and includes data from – a range of stakeholders and different viewpoints throughout the evaluation

- Is open and flexible in terms of changing the design and data collected, in response to changes and new information emerging as the intervention is delivered
- Has incorporated, in a realistic way, evaluation approaches which may incorporate different underpinning theories of causality
- Transparency about all the above
- Contributory in advancing wider knowledge or understanding
- Defensible in design by providing an evaluation strategy that can address the evaluation questions posed
- Rigorous in conduct through the systematic and transparent collection, analysis and interpretation of data
- Credible through offering well-founded and plausible arguments about the significance of the data generated
- The scope for drawing wider inference is explained
- The evaluation design is defensible and is suitable for a complex policy
- The contexts of data sources are retained and portrayed
- The outcome of the evaluation is plausible, feasible, appropriate in scope and has organisational “fit”, that is the recommendations are usable within the organisation/ programme

Examples of established quality standards are in Chapter 4 **Appendix 1: Tools and resources**.

## 4.4 Data Quality

There are some generally recognised and common challenges to data collection that need to be addressed which include the following

- Data collection needs to be systematic and built into the evaluation project: it needs to be collected at regular, scheduled intervals
- Effective data collection and analysis starts with clarity of purpose – what data needs to be collected, what questions the data will answer and how the answers will contribute to the evaluation and future programme decisions
- Methods of data analysis need to be agreed and effectively utilised and also undertaken at regular intervals so that trends, data gaps etc., can be detected
- Results of the analysis must be shared with stakeholders as part of the regular review point agenda
- Data access, storage and data sharing and ethical agreements needs to be agreed and signed off in advance

### 4.4.1 The challenges of “big data”

Big data refers to data sets that are complex and large and which cannot be analysed by traditional statistical tools. Large amounts of data are now routinely gathered and stored as a consequence of the exponential growth of digital devices and networks across the world. This data can be analysed to identify trends and patterns in behaviour and used to inform decision-making.

Big data is increasingly informing evaluations. This presents its own challenges:

- Volume – the sheer amount of data that is now routinely collected
- Velocity – the speed with which the data is gathered and updated
- Variety – this refers to the range of detail that current big data gathering systems can collect

Faced with these challenges, the same actions noted in section 4.3 above apply:

- The information gathered needs to be appropriate to purpose
- Means of shaping the data for analysis need to be agreed including storage and linkage to other datasets

## Chapter 5: Achieving impact

As earlier chapters emphasised, evaluations are conducted for a purpose. This chapter considers how one can help to ensure that the intended purpose is fulfilled. The first section (5.1) discusses what happens after a complex evaluation and why this matters. Section 5.2 examines some of the barriers and enablers to the implementation of evaluation outcomes. Section 5.3 considers the use of past evaluations as an often-overlooked resource. Finally, section 5.4 discusses the need for increasing the capacity and capabilities of evaluation staff. The key message of this chapter is the importance of embedding the evaluation process into the overall work of the organisation.

For more information about the issues and any of the resources discussed in this chapter, see the corresponding tools and resources for Chapter 5 in **Appendix 1: Tools and resources**.

### 5.1 What happens after a complex evaluation, and why does it matter?

In general, the use of evaluation findings is taken to mean that the factors that impact on the outcomes of a policy or programme are understood and inform its subsequent direction and resourcing – so called instrumental usage.

However, other sorts of evaluation use are also important (Weiss, 1998) such as insights from learning and shifts in thinking over time (conceptual uses).

Even strategic uses of evaluation – to defend a position or resist change to a programme – can provide useful signals about its importance, for example in drawing attention to a lack of commitment to change in an organisation. Hence, many aspects of an evaluation can be used and there are many different potential users of its findings, from programme managers or policy analysts, to wider stakeholders and even civil society (Weiss, 1998).

Understanding the impact of a policy or programme and ensuring evaluation findings feed back into the policy process is vital and evaluation needs to be integrated into a continuous process of adaptation, learning and policy evolution that accepts the challenges and uncertainties inherent in complexity as a requirement of good policy making.

### 5.2 What are the barriers/enablers to complex evaluations being used and having impact?

#### 5.2.1 Commitment to learning and change

Organisational/institutional buy-in and commitment to complexity approaches are essential to help create the right conditions for complexity approaches to policy making, including acceptance of uncertainty, political will and commitment to learning. These conditions can help evaluation uptake; issues that require longer-term evaluations require additional commitment and political will from stakeholders.

Defra's **Complexity Evaluation Framework** (CEF) (2019) is a good example of a learning resource that was commissioned to help embed complexity thinking into evaluation practice. It emphasises the importance of the regular feedback and embedding of learning from evaluation into policy design when working in complex policy settings.

The CEF is part of a wider approach to complex systems thinking in the Department, led by the Chief Scientific Advisor's Office (Oliver, 2020).

### 5.2.2 Making complexity actionable can help with uptake

Embracing complexity-appropriate approaches in evaluation practice can be similarly challenging - there may be trade-offs between complexity approaches and the resources needed to undertake complexity-appropriate evaluation with the perception that complexity approaches are 'nice to have' but difficult to do in practice, with uncertainty over longer term impacts of a policy or programme. This in turn may limit the perceived value of evaluation findings. Hence turning 'overwhelming complexity' into 'actionable complexity' through practical and useful ways that engage stakeholders is critical to success and can help with evaluation uptake. For an example of one such approach, see Barbrook-Johnson and Penn (2021).

### 5.2.3 Communicating findings is key

The *Magenta Book Supplementary Guidance* (2020) acknowledges that complexity is difficult to communicate and may need additional time and resources to help stakeholders understand the findings of an evaluation.

It highlights evaluation approaches that can help engage stakeholders in ways that are meaningful to them and hence assist with evaluation dissemination and uptake (see bullet points below). In addition, the choice of methodology, building in specific plans for communication at the design stage of an evaluation is important, even if those plans have to be flexed and adapted as an evaluation proceeds.

CECAN Evaluation and Policy Practice Note (No.8) *Maximising impact from evaluations in complex policy areas* contains a useful communications checklist for helping ensure uptake and impact from evaluations (Scott, 2017):

- Write dissemination plans into the design from the outset and factor in resources for drawing on the expertise of the consultants in designing and carrying these out
- Plan to disseminate widely both internally and further afield as appropriate
- Tailor the messages and media for different audiences
- Highlight interesting, important and engaging findings to build enthusiasm
- Use a flexible approach and seize communication opportunities
- Integrate the communication with other initiatives to build value and impact
- Have plans in place for communicating where things didn't work and gaining valuable learning from this information
- Build connections with other policy leads who want to apply the lessons in their own areas



The *Evaluation Toolkit* (The Pell Institute et al, 2014) has a useful set of questions that can be used to develop a communications plan for the dissemination of evaluation results.

The *Magenta Book Supplementary Guidance* (CECAN 2020) explains how different types of methods and approaches can help with the communication challenges. These are summarised in Figure 7 below.

Useful methods and approaches	How they help with communication challenges
Participatory approaches	Generate deeper, shared understanding and trust among those involved
Narrative methods	Help people engage effectively with the evaluation and communicate findings relating to the impacts on people and communities more meaningfully
Agent based modelling	Provides a method of generating narratives that explain results in ways people can relate to

**Figure 7:** Useful methods and approaches for tackling communication challenges arising from complexit. Source: CECAN (2020)

The nature of the communications is also important: how do you create simplified products and outputs from complex evaluations? End products need to be engaging and usable, so it is worth considering a hierarchy of products from simple infographics through to raw data tables, bearing in mind the communication channel and audiences. Infographics and other visual techniques such as video and animation also lend themselves well to social media platforms. Examples of effective means of communicating complexity visually are available in Chapter 5 **Appendix 1: Tools and resources**.

**5.2.4 Timing it right can make a difference**

Timeliness of evaluation findings, and findings that are ‘good enough’ can help evaluation findings feed back into policy. There is often the greatest interest and enthusiasm around a policy at its inception and yet evaluation often comes on-line much later in the policy cycle. Producing evaluation findings early on in the process can be helpful and engage key audiences.

Giorgi (2017) provides useful insights here, based on learning from an evaluation of Defra’s Reward and Recognition Fund.

If evaluators are encouraged to share emerging findings and insights at a point in time as ‘good enough’ and policy makers, in turn, accept the risk that final results and end conclusions may differ, then the evaluation’s potential to achieve impact is greatly enhanced through a continuous cycle of learning, rather than leaving feedback until

the end of an evaluation when interest in findings has passed and priorities may have changed.

“Evaluation needs to be an integrative, continuous process not a one-off exercise at the end or a series of self-contained steps, it needs to become a way of working” (Giorgi, 2017)

### 5.3 Analysing past evaluations: an overlooked resource

A major gap in policy making is learning the lessons from past interventions and integrating the lessons from evaluations that have been undertaken. Producing a meta evaluation – literally an evaluation of evaluations – is another way of trying to ensure findings from complex evaluations are used, and useful.

Using a mixed methods approach, a CECAN partner carried out a meta-evaluation of 23 evaluations in nexus policy domains (Collingwood Environmental Planning Ltd, 2016). The research aimed to learn the lessons from past policy evaluations and understand the factors that support or inhibit successful evaluations, including whether an evaluation met its objectives and the impacts or use that the evaluation had (instrumental, conceptual, strategic or process).

### 5.4 Commitment to building capacity and capability in evaluation staff

As mentioned above, failure to invest in learning and capacity for complex evaluation can be a barrier to uptake. Good complex evaluation can only take place when there is the capacity – and individual capabilities – to understand and address the evaluation challenges posed by complexity.

It is generally accepted that there are three levels in building evaluation capacity

- Cultural
- Organisational
- Individual

From a complex dynamic systems perspective, it is understood that systems operate at multiple levels and changes taking place at one level influence what happens at the levels above and below. For example, providing guidance and training to an individual analyst in complex evaluation approaches will not substantially change practice, if the overall commissioning practices deter a ‘complexity appropriate’ approach. On the other hand, attempts to develop an overall ‘complexity sensitive’ culture in an organisation will also require the provision of training and guidance to individuals in the use of appropriate evaluation designs.

#### 5.4.1 Complex evaluation capabilities

Although there are now a number of capability and competency frameworks available that outline the skills and qualities required for high quality evaluation,

there is currently no capability framework developed specifically for evaluators and commissioners working in complex settings. There are, however, capability frameworks that address areas that are particularly relevant to complex settings.

We focus on two which are particularly relevant in the UK context: The UK **Government analytical evaluation capabilities framework** (HM Treasury, 2020b) and the **UK Evaluation Society's Framework of Evaluation Capabilities** (2012).

The UK *Government analytical evaluation capabilities framework* (HM Treasury, 2020b) sets out the skills, attitudes and practices which enable effective and high-quality Government evaluations. It has a particular focus on the capabilities required at different stages of planning, commissioning and managing an evaluation.

Particularly pertinent to complexity are the following capacities

- Ensures evaluation is considered early in the *ROAMEF policy cycle* and that enough thought is given to monitoring and evaluation requirements (including resources needed) at the appraisal stage e.g., in the business case (*this is a message that also runs throughout this Toolkit*)
- Demonstrates knowledge and experience of a range of types and approaches to evaluating programmes and policies (*particularly important to have an understanding of complex evaluation approaches*)

In the 'understanding the intervention and constructing a *Theory of Change* section:

- Can create a *Theory of Change*, linking elements of the intervention with outcomes (*recommended in many complex evaluation designs*)

Identifying the right evaluation design for the policy (*i.e., using a complex appropriate evaluation approach if the situation requires this*)

- Understands the range of approaches which can be used to assess whether an intervention was effective in meeting its objectives, and can determine the most appropriate approach and methods for a particular evaluation (*particularly understanding of the range of approaches relevant for complex evaluations*)
- Is aware of frameworks for assessing the complexity of a programme and the distinctions between: 'simple; complicated and complex'

The *Evaluation Society's Framework of Evaluation Capabilities* (2012) has a stronger focus on the values and ethics of evaluation practice than on the technical skills required. It addresses a number of capabilities that are of particular relevance to complex evaluations.

In section 1 on evaluation knowledge:

- 1.21 Is familiar with a range of evaluation theories and approaches (*including approaches suitable for complex evaluations*)
- 1.24 Knows how to design an evaluation appropriate to the evaluation task
- 1.25 Takes account of the policy context in designing evaluations (*context is particularly important in complex adaptive systems*)

1.35 Comprehends that criteria for validity and reliability differ according to method (*this is particularly pertinent to the different principles of causality and attribution that underpin some of the 'complex appropriate' evaluation approaches*)

In section 2 on professional practice:

2.11 Ascertains the social /political context and programme logic

2.14 Gathers relevant evidence, analyses and interprets in context (*see above point on importance of context in complex adaptive systems*)

2.22 Shows ethical sensitivity in specific socio/political contexts (ditto)

In section 3 on qualities and dispositions:

3.1 Demonstrates ability to adapt to changing circumstances in a principled manner (*Complex adaptive systems are constantly in change, evaluation methods sometimes have to be adapted in response to this*)

3.4 Displays independence of mind and integrity especially when evaluation challenged (*Evaluators may find themselves being challenged by stakeholders unfamiliar with the evaluation approaches they are using*)

In addition, there are a number of CECAN resources which can be used for developing individual capacity for complex evaluation. These are available under the resources for Chapter 5 in **Appendix 1: Tools and resources**.

# Appendix 1: Tools and resources

## Chapter one

1. CECAN (2020). *Magenta Book 2020 Supplementary Guide: Handling Complexity in Policy Evaluation*. HM Treasury. Available at: <https://www.gov.uk/government/publications/the-magenta-book>
2. CECAN Ltd., Risk Solutions & Tavistock Institute. (2019). *Complexity Evaluation Framework*. Department for Environment, Food and Rural Affairs (Defra). Available at: <http://sciencesearch.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Completed=220&ProjectID=20401>  
Commissioned by Defra, the *Complexity Evaluation Framework* is a framework of key considerations to guide the scoping, commissioning, management and delivery of complexity-appropriate evaluations.

## Chapter two

3. Cox, J., & Barbrook-Johnson, P. (2020). How does the commissioning process hinder the uptake of complexity-appropriate evaluation? *Evaluation*, 27(1), 32–56. <https://doi.org/10.1177/1356389020976157>  
This article discusses the key points in chapter 2 in more detail.
4. Cox, J. (2019). *How does the commissioning process inhibit the uptake of complexity-appropriate evaluation?* Retrieved from <https://www.cecan.ac.uk/wp-content/uploads/2020/09/Jayne-Cox-Report-final.pdf>  
This report discusses the key points in chapter 2 in more detail.
5. BetterEvaluation (2013) 'Manage an evaluation or evaluation system', in *The Rainbow Framework*. Available at: [https://www.betterevaluation.org/en/rainbow\\_framework/manage\\_evaluation](https://www.betterevaluation.org/en/rainbow_framework/manage_evaluation)  
BetterEvaluation's Rainbow Framework is a useful guide to the different steps and considerations in monitoring and evaluation. The 'manage an evaluation' segment lists a number of tasks involved in managing an evaluation and suggests a number of methods and processes for doing these, including deciding who will conduct the evaluation and who will make decisions about it.
6. Macmillan, T. (2020) *Experiences scoring and assessing complex evaluations*. Retrieved from: <https://www.cecan.ac.uk/wp-content/uploads/2020/09/Tarran-Macmillan-Defra1.pdf>
7. BetterEvaluation. (2018). *Manager's guide to evaluation*. Available at: [https://www.betterevaluation.org/commissioners\\_guide](https://www.betterevaluation.org/commissioners_guide)  
BetterEvaluation's comprehensive and interactive guide to commissioning and managing an evaluation.
8. University of Sheffield, & CFE Research. (2014). *Good practice guide: commissioning evaluations*. Available at: <https://golab.bsg.ox.ac.uk/knowledge-bank/resources/good-practice-guide-commissioning-evaluations/>  
A useful guide to commissioning an independent evaluation and some of the issues to be considered, written with the aim of supporting grant-supported multi-agency partnerships who must commission an evaluation in order to

access the grant.

9. Centre for Epidemiology and Evidence. (2019). *Commissioning Evaluation Services: A Guide*. 2nd ed. Evidence and Evaluation Guidance Series, Population and Public Health Division. Sydney: NSW Ministry of Health. Retrieved from: <https://www.health.nsw.gov.au/research/Publications/evaluation-guide.pdf>

Another comprehensive guide to commissioning evaluations developed to support New South Wales Health staff in the commissioning of population health program evaluations.

10. European Monitoring Centre for Drugs and Drug Addiction. (2017). *Evaluating drug policy: a seven-step guide to support the commissioning and managing of evaluations*. Publications Office of the European Union, Luxembourg. Retrieved from: [www.emcdda.europa.eu/system/files/publications/4680/td0417390enn.pdf\\_en](http://www.emcdda.europa.eu/system/files/publications/4680/td0417390enn.pdf_en)

This tool summarises the main issues that commissioners of evaluations need to consider along with links to further sources of information and advice.

11. McNeish, D. & Scott, S. (2016). *DMSS guide to project evaluation: Part 3 Commissioning an independent evaluation*. Retrieved from: <https://www.dmss.co.uk/pdfs/Part-3-Commissioning-an-independent-evaluation.pdf>

Further practical guidance on commissioning an evaluation by an organisation that undertakes research and evaluations in the social care sector.

12. CECAN (2020). *Magenta Book 2020 Supplementary Guide: Handling Complexity in Policy Evaluation*. HM Treasury. Available at: <https://www.gov.uk/government/publications/the-magenta-book>

Chapter 3 of this Magenta Book Supplementary Guide is particularly relevant to those commissioning a complex evaluation.

## Chapter three

13. Stabilisation Unit. (2020). *Monitoring, Evaluation and Learning (MEL) in Conflict and Stabilisation Settings: A Guidance Note*. Available at: <https://www.gov.uk/government/publications/monitoring-evaluation-and-learning-mel-in-conflict-and-stabilisation-settings-a-guidance-note>

An HM Government guidance note, for use in conflict and stabilisation settings but with wider applicability in setting up a MEL framework.

14. BetterEvaluation (2013) 'Manage an evaluation or evaluation system', in *The Rainbow Framework*. Available at: [https://www.betterevaluation.org/en/rainbow\\_framework/manage\\_evaluation](https://www.betterevaluation.org/en/rainbow_framework/manage_evaluation)

This covers the key points in setting up, designing and managing an evaluation system.

15. OECD-DAC. (2010). *Glossary of key terms in evaluation and results based management*. Paris: OECD-DAC. Retrieved from: <http://www.oecd.org/development/peer-reviews/2754804.pdf>

16. BetterEvaluation. (2014). *Evaluability Assessment*. Available at: [https://www.betterevaluation.org/en/themes/evaluability\\_assessment](https://www.betterevaluation.org/en/themes/evaluability_assessment)

A useful guide to evaluability assessments.



17. CECAN (2020). *Magenta Book 2020 Supplementary Guide: Handling Complexity in Policy Evaluation*. HM Treasury. Available at: <https://www.gov.uk/government/publications/the-magenta-book>  
Chapter 4 of the Supplementary Guide is particularly relevant. It provides further background information on complexity appropriate evaluation strategies.
18. HM Treasury. (2020a). *The Magenta Book*. Available at: <https://www.gov.uk/government/publications/the-magenta-book>  
This is the UK Central Government guidance on evaluation.
19. Rogers, P. (2014). *Theory of Change: Methodological Briefs - Impact Evaluation No. 2*. Retrieved from <https://www.unicef-irc.org/publications/747-theory-of-change-methodological-briefs-impact-evaluation-no-2.html>  
This document provides more detail about the role of Theory of Change (ToC) in evaluation.
20. Wilkinson, H., Hills, D., Penn, A., & Barbrook-Johnson, P. (2021). Building a system-based Theory of Change using Participatory Systems Mapping. *Evaluation*, 27(1), 80–101. <https://doi.org/10.1177/1356389020980493>  
This paper describes a process for building a system-based Theory of Change (ToC).
21. Penn, A., & Barbrook-Johnson, P. (2020). *Participatory Systems Mapping: a practical guide*. Retrieved from <https://www.cecan.ac.uk/wp-content/uploads/2020/09/PSM-Workshop-method.pdf>  
This document describes a process for building a system map with stakeholders.
22. Befani, B. (2020). *Choosing Appropriate Evaluation Methods – A Tool for Assessment and Selection*. Version 2. Available at: <https://www.cecan.ac.uk/news/choosing-appropriate-evaluation-methods-a-tool-for-assessment-and-selection-version-two/>  
The Choosing Appropriate Evaluation Methods tool was originally developed by Barbara Befani and Michael O'Donnell. Version 2 of the Excel-based tool and its accompanying report have been expanded and partially redeveloped to include more complexity-appropriate methodologies. This is the CECAN sponsored tool to assist in choosing potential evaluation methods.
23. Castellani, B. C., & Rajaram, R. (2021). *Data Mining and Big Data*. SAGE Quantitative Research Kit (Volume 11). SAGE Publications. Available at: <https://uk.sagepub.com/en-gb/eur/big-data-mining-and-complexity/book257490>

## Chapter four

24. CECAN (2020). *Magenta Book 2020 Supplementary Guide: Handling Complexity in Policy Evaluation*. HM Treasury. Available at: <https://www.gov.uk/government/publications/the-magenta-book>  
Chapter 3 of the Supplementary Guide is particularly relevant to those managing a complex evaluation.
25. CECAN Ltd., Risk Solutions & Tavistock Institute. (2019). *Complexity Evaluation Framework*. Department for Environment, Food and Rural Affairs (Defra). Available at <http://sciencesearch.defra.gov.uk/Default.aspx?Menu=Menu&>

**Module=More&Location=None&Completed=220&ProjectID=20401**

Commissioned by Defra, the *Complexity Evaluation Framework* is a framework of key considerations to guide the scoping, commissioning, management and delivery of complexity-appropriate evaluations. This framework is written for commissioners and managers of evaluations. Considerations for managing complexity-appropriate evaluations are embedded throughout each of the three core chapters of the framework: understanding, designing and embedding.

26. O'Donnell, M. (2016). *Adaptive management: what it means for CSOs*. Bond. Available at: <https://www.bond.org.uk/resources/adaptive-management-what-it-means-for-csos>  
This report discusses adaptive management - an approach to working on complex problems or contexts, where solutions cannot be completely known in advance and, therefore, that interventions cannot be planned out in full ahead of time.
27. Puttick, R. (2018). *Mapping the Standards of Evidence used in UK social policy*. Alliance for Useful Evidence. Available at: <https://www.alliance4usefulevidence.org/publication/mapping-the-standards-of-evidence-used-in-uk-social-policy/>  
A review of some of the recently developed standards of evidence in evaluations.
28. Spencer, L., Ritchie, J., Lewis, J., & Dillon, L. (2004). *Quality in qualitative evaluation: a framework for assessing research evidence*. Government Chief Social Researcher's Office. Available from: <https://www.gov.uk/government/publications/government-social-research-framework-for-assessing-research-evidence>  
This report offers a framework to guide assessments of the quality of qualitative research evaluations.
29. UK Evaluation Society. (2018). *Guidelines for Good Practice in Evaluation*. Available at: <https://www.evaluation.org.uk/professional-development/good-practice-guideline/>  
These guidelines are designed to help commissioners, practitioners and participants establish good practice in evaluation. The report sets out both technical and ethical standards, including around equity and diversity.
30. Taplin, D., Clark, H., Collins, E., & Colby, D. (2013). A Good Quality Theory of Change. In *Theory of Change Technical Papers*. ActKnowledge. Retrieved from <https://www.actknowledge.org/resources/documents/ToC-Tech-Papers.pdf>  
Section 4 of this report sets out the features of a good quality Theory of Change (ToC).
31. The RAMESSES II Project. (2017). *Quality Standards for Realist Evaluation: For Evaluators and Peer Reviewers*. Available at: [https://www.ramesesproject.org/Standards\\_and\\_Training\\_materials.php](https://www.ramesesproject.org/Standards_and_Training_materials.php)  
This resource sets out quality standards for evaluators and peer-reviewers of realist evaluations.
32. CECAN (2020). *Magenta Book 2020 Supplementary Guide: Handling Complexity in Policy Evaluation*. HM Treasury. Available at: <https://www.gov>

**[uk/government/publications/the-magenta-book](https://www.gov.uk/government/publications/the-magenta-book)**

For further information and helpful tables for use in planning complexity-appropriate evaluation approaches, see chapter 4 of the Supplementary Guide.

## Chapter five

33. Weiss, C. H. (1998). Have We Learned Anything New About the Use of Evaluation? *American Journal of Evaluation*, 19(1), 21–33. <https://doi.org/10.1177/109821409801900103>
34. CECAN Ltd., Risk Solutions & Tavistock Institute. (2019). *Complexity Evaluation Framework*. Department for Environment, Food and Rural Affairs (Defra). Available at: <http://sciencesearch.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Completed=220&ProjectID=20401>  
Commissioned by Defra, the *Complexity Evaluation Framework* (CEF) (2019) is a framework of key considerations to guide the scoping, commissioning, management and delivery of complexity-appropriate evaluations. Section 6 of the CEF explores the implications of complexity for ‘embedding’. This is the element of an evaluation concerned with feeding back understanding and learning to evaluation users and participants and embedding these into relevant processes both inside and outside of the evaluation (e.g., dissemination and use).
35. Oliver, T. (2020). *A Systems Approach to Environmental Policy in Defra*. CECAN Webinar. Available at: <https://www.cecan.ac.uk/events/cecan-webinar-a-systems-approach-to-environmental-policy-in-defra/>  
This webinar recording discusses systems approaches to environmental policy in Defra.
36. Barbrook-Johnson, P., & Penn, A. (2021). Participatory systems mapping for complex energy policy evaluation. *Evaluation*, 27(1), 57–79. <https://doi.org/10.1177/1356389020976153>  
This paper presents Participatory Systems Mapping and how it can be of value in evaluation.
37. Scott, K. (2017). *Maximising impact from evaluations in complex policy areas*. CECAN Evaluation and Policy Practice Note (No. 8). Available at: <https://www.cecan.ac.uk/resources/eppns/>  
This CECAN Evaluation and Policy Practice Note contains a useful communications checklist to assist with dissemination and uptake of evaluations.
38. The Pell Institute, IHEP, and Pathways to College Network. (2014). Develop a communications plan. *The Evaluation Toolkit*. Retrieved from: <http://toolkit.pellinstitute.org/evaluation-guide/communicate-improve/develop-a-communications-plan/>  
This section of the Evaluation Toolkit website outlines a set of questions that can be used to develop a communications plan for the dissemination of evaluation results.
39. Boehnert, J., Penn, A., Barbrook-Johnson, P., Bicket, M., Hills, D., (2018). *The Visual Representation of Complexity – Definitions, Examples and Learning Points*.

Retrieved from <https://www.cecan.ac.uk/sites/default/files/2018-07/JB%20online%20pdf%20The%20Visual%20Communication%20of%20Complexity%20-%20May2018%20-%20EcoLabs.pdf>

This poster seeks to communicate sixteen key characteristics of complexity visually.

40. Gunton, R. (2018). *The Pluralistic Evaluation Framework*. Retrieved from <https://www.cecan.ac.uk/wp-content/uploads/2020/08/Poster-04-Richard-Gunton.pdf>

This poster is an example of using visual methods to communicate evaluation findings.

41. Giorgi, S. (2017). *How to improve the evaluation of complex systems to better inform policymaking – Learning from evaluating Defra’s Reward & Recognition Fund*. CECAN Fellowship Report. Retrieved from: <https://www.cecan.ac.uk/sites/default/files/2018-01/Guidance%20Report%20-%20RRF%20Fellowship%20Final.pdf>

Using the evaluation of Defra’s Reward and Recognition Fund as a case study, this report presents insights into policy and practitioner’s evaluation experiences and challenges. It offers some particularly valuable conclusions into how to improve the dynamics between evaluation and policymaking regarding the policy cycle when in complex settings.

42. Collingwood Environmental Planning Ltd (2016). *Learning lessons for evaluating complexity at the nexus: A meta-evaluation of CEP projects*. CECAN Report. Retrieved from: <https://www.cecan.ac.uk/wp-content/uploads/2020/09/Learning-lessons-for-evaluating-complexity-at-the-nexus-CEP-FINAL-070317-for-publishing.pdf>

This report highlights research that aimed to learn the lessons from past policy evaluations and understand the factors that support or inhibit successful evaluations, including whether an evaluation met its objectives and the impacts or use that the evaluation had had (instrumental, conceptual, strategic or process). A summary of this report is also available as a CECAN Evaluation and Policy Practice Note from: <https://www.cecan.ac.uk/wp-content/uploads/2020/08/EPPN-No-05-Learning-Lessons.pdf>.

43. HM Treasury (2020b). *Government analytical evaluation capabilities framework*. Available at: <https://www.gov.uk/government/publications/the-magenta-book>

This supplement to the Magenta Book brings together and summarises the knowledge and skills required by evaluation managers to design and deliver evaluations in government. It complements the corresponding chapter in the Magenta Book (chapter 7 on evaluation capabilities) with further information and a self-assessment tool for evaluation managers to use.

44. UK Evaluation Society. (2012). *Framework of Evaluation Capabilities*. Available at: <https://www.evaluation.org.uk/professional-development/framework-of-evaluation-capabilities/>

This capabilities framework can be used by individuals, commissioners and/or organisations in training and professional development and as a form of quality assurance. It provides an opportunity to reflect upon and strengthen whatever

capabilities they consider necessary to enhance their practice. Relative to the Government analytical evaluation capabilities framework above, it has a stronger focus on the values and ethics of evaluation practice required.

45. CECAN (2017). *Evaluation of Complex Policy and Programmes - A CECAN module for future policy analysts and evaluators*. November 2017, Version 1.0. Retrieved from: [https://www.cecان.ac.uk/wp-content/uploads/2020/09/Cecan-Module-Syllabus\\_17-Dec.pdf](https://www.cecان.ac.uk/wp-content/uploads/2020/09/Cecan-Module-Syllabus_17-Dec.pdf)

This is a CECAN Module for future policy analysts and evaluators. The course was developed to support capacity building nationally and internationally in the evaluation of complex policy and programmes.

46. CECAN (multiple dates). *CECAN Evaluation and Policy Practice Notes (EPPNs)*. Available at: <https://www.cecان.ac.uk/resources/eppns/>
47. CECAN (n.d.). *CPD Courses*. Available at: <https://www.cecان.ac.uk/training/>
48. CECAN (multiple dates). *Videos*. Available at: <https://www.cecان.ac.uk/videos/>
49. CECAN (multiple dates). *Resources*. Available at: <https://www.cecان.ac.uk/resources/>

CECAN produces a variety of resources available for download. They are all available under a Creative Commons CC BY 4.0 licence. In addition to the CECAN EPPNs and course module syllabus linked above, further resources available include posters, conference papers, project reports, journal articles and toolkits, among others.

## Appendix 2: Glossary of terms and acronyms

### **Agent based modelling:**

Agent-based models seek to demonstrate the emergent behaviour and properties of systems by modelling the behaviours of the many individual actors that make up the system, the agents, and the interactions between them. Agents can be people, companies, projects, assets, vehicles, cities, animals, etc. Agent-based modelling differs from many classical modelling approaches in that: No attempt is made to impose the behaviour of the system directly; the global behaviour emerges as a result of interactions of many individual behaviours. There is no assumption of equilibrium; the system is modelled as dynamic and adaptive.

### **Approach:**

The approach refers to the overall, conceptual perspective adopted for the evaluation – for example, a theory-based approach articulates and tests a theory of how the policy causes or contributed to observed results. Different approaches can be combined in a hybrid design.

### **Confounding factors:**

These are variables which may not have been initially taken into account in the evaluation and yet have a significant impact on the system's functioning and outcomes. Complex systems potentially have many confounding factors, and it is important to identify these at an early stage of evaluation design.

### **Counterfactual approaches:**

Counterfactual approaches are interested in three key issues: (1) proof that, in the absence of some policy intervention, the outcome would have stayed the same; (2) identification of instances where the policy intervention did not produce the desired results, either as predicted or to the extent expected, or produced undesired or unexpected results (as in the case of increasing health harms or inequalities); (3) or, finally, evidence that, contrary to expectation, a desired result took place but was not a function of the policy intervention.

### **Emergence:**

New, unexpected higher-level properties can arise from the interaction of components. These properties are said to be emergent if they cannot easily be described, explained, or predicted from the properties of the lower-level components.

### **Feedback loops:**

Feedback occurs when outputs of a system are routed back as inputs as part of a chain of cause-and-effect that forms a circuit or loop. The system can then be said to feed back into itself.

### **Generative causation:**

An approach which seeks to articulate the underlying mechanisms or processes of change, and test the theory empirically to investigate whether, why or how the policy causes or contributed to observed results, and how context influences these.



**Levers and hubs:**

These are components of a system that have a disproportionate influence because of the structure of their connections. How these behave can help to mobilise change, but their behaviour may also make a system vulnerable to disruption.

**Method:**

A method is a way of gathering, analysing or making sense of data. Methods can be, but are not necessarily, independent of the approach adopted. Examples of data collection methods include surveys, questionnaires, interviews, realist interviews (specific to realist evaluation), desk reviews, and (critical) observation.

**Participatory systems mapping:**

A systems map is a graphical representation of the components in a system and the causal relationships between them. Participatory systems mapping is a participative approach for building and analysing systems maps; it can help to build a common understanding of system complexity across different stakeholder communities and enable a structured means of communicating the learning around these issues. Participatory systems mapping is part of a wider family of approaches which includes causal loop analysis, systems dynamics, conceptual mapping and group model building.

**Qualitative Comparative Analysis:**

Qualitative Comparative Analysis (QCA) is a means of analysing the causal contribution of different conditions (e.g., aspects of an intervention and the wider context) to an outcome of interest. QCA starts with the documentation of the different configurations of conditions associated with each case of an observed outcome. These are then subject to a minimisation procedure that identifies the simplest set of conditions that can account all the observed outcomes, as well as their absence.

**ROAMEF policy cycle:**

The Rationale, Objectives, Appraisal, Monitoring, Evaluation, Feedback cycle is promoted by the UK Government to ensure policy makers receive evidence of whether change programmes are achieving their aims and objectives. It is a proven, sequential, robust evaluation process. It is systematic and follows a logical process.

**Sensitivity to context:**

From the perspective of evaluation, findings which are sensitive to the social, political, ecological and temporal context. This means that generalizations across the system or outside it should be treated with caution.

**Strategy:**

An evaluation strategy refers to the whole evaluation plan, including the choice of approach, methods and tools, and how these will be put in place.

**Systems mapping and modelling:**

See Participatory system mapping.

**Theory of Change (ToC):**

A Theory of Change (ToC) explains how the programme/intervention is understood to produce a series of outputs that contribute to achieving the intended impacts. In an evaluation, a Theory of Change is useful for identifying the data to be gathered and analysed. It can also provide a basis for reporting the findings of the evaluation.

## Acronyms

CECAN: Centre for the Evaluation of Complexity Across the Nexus

CEF: Complexity Evaluation Framework

Defra: UK Department for Environment, Food and Rural Affairs

DiD: Difference in Difference

ITT: Invitation to tender

MEL: Monitoring, Evaluation and Learning framework

NGO: Non-Governmental Organisation

PSM: Participatory System Mapping

QCA: Qualitative Comparative Analysis

RFP: Request for proposals

ROAMEF: Rationale, Objectives, Appraisals, Monitoring, Evaluation and Feedback

ToC: Theory of Change

**CECAN is working with following organisations and many others**



## **Our Partners and associates**

