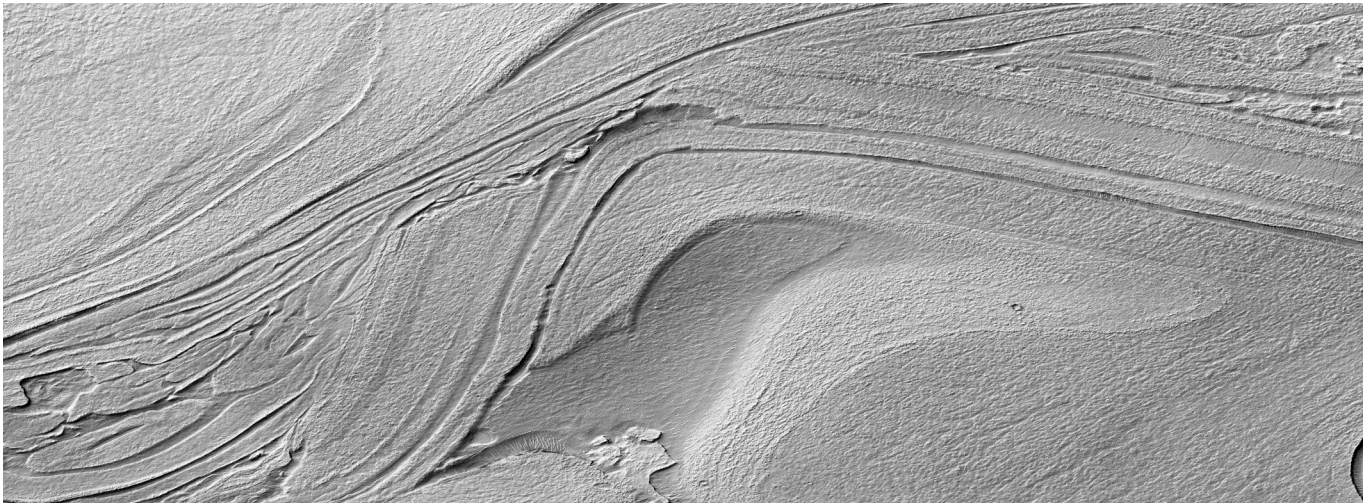# Learning lessons for evaluating complexity across the nexus: A meta-evaluation of projects

A CECAN Evaluation and Policy Practice Note for policy analysts and evaluators

A major gap in policy making is learning the lessons from past interventions and integrating the lessons from evaluations that have been undertaken. A meta-evaluation, an evaluation of evaluation studies, was undertaken by Collingwood Environmental Planning Ltd (CEP) as an intensive piece of research (July-November 2016) to provide early outputs as part of CECAN's scoping process. A sample of 23 projects (out of a total of 43) was selected on the basis of publicly available documentation (final evaluation reports) and current CEP staff with knowledge of those projects from across the nexus issues, including evaluations of flood risk, biodiversity, landscape, land use, climate change, catchment management, community resilience, bioenergy, and EU Directives.
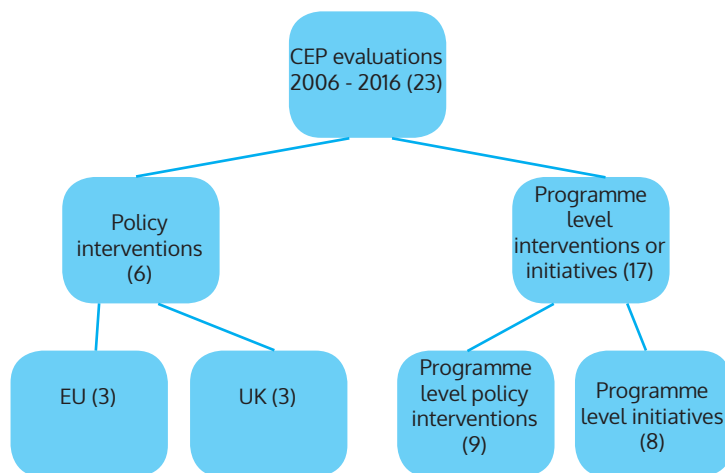
What was the purpose of the meta-study?

The aims were:
- To learn the lessons from past policy evaluations (using CEP projects)
- To understand the factors that support or inhibit (barriers or enablers to) successful evaluations, where success is measured by:
  - » Whether the evaluation meets its own objectives
  - » The impact (or use) that the evaluation has in four categories:
    - Instrumental – evidence has a direct impact on policy
    - Conceptual – evidence influences how stakeholders think about a policy area/issue
    - Strategic – evidence used for accountability and defending/promoting policy
    - Process – improved working processes in some way
- To explore the value of different types of approaches and methods used for evaluating complexity.

What approach was adopted?

A multiple embedded case study design was adopted, selecting 23 cases which were categorised as follows (the numbers indicate number of cases in each category).

CEP evaluations 2006 - 2016 (23)

- Policy interventions (6)
  - EU (3)
  - UK (3)
- Programme level interventions or initiatives (17)
  - Programme level policy interventions (9)
  - Programme level initiatives (8)

These categories were based on common policy contexts in which the evaluations were taking place, rather than, for example, the type of evaluation (e.g. formative, summative) or the specific nexus sector since many of the cases covered multiple types of evaluation and sector. Each case was then categorised in a master Excel spreadsheet according to the following criteria: Scale; Theme; Evaluation type; Evidence collection method; Type of complexity; and Type of evaluation use.

Four key Meta-Evaluation Questions provided the evaluation framework and structure:
- Were the evaluations fit for purpose, and was their purpose clear?
- Has the framing of the evaluation been more or less useful for understanding complexity?
- What methods have been used for dealing with aspects of complexity found within environmental policy and which methods appear to have been most effective?
- What factors lead to an evaluation being more (or less) influential on policy changes/ outcomes / evaluation use?

These were used to structure more specific evaluation questions, answered on the basis of documented project final reports and supplemented by interviews with CEP project managers.

What were the key findings?

In the case of Fitness for Purpose:
- In general the evaluations were fit for purpose, but often because they were tailor made to the circumstances.
- It is important to ensure objectives of a policy/intervention are linked to a clear baseline and that there are specific measurable outcomes that an evaluation can then assess.
- Scope to discuss, amend and agree evaluation objectives as part of the initial work on an evaluation helps ensure clarity and fitness for purpose, and ongoing reflection on evaluation objectives is important especially when the policy objectives may be evolving during the course of the project.
- Exploratory interventions focused on learning and process can create uncertainty in what is to be delivered and why. More attention from stakeholders is required to avoid vaguely defined/inconsistent policy/intervention objectives.
- Setting clear programme (higher) level objectives at the outset to reflect the relationship between the programme and individual project level (e.g. local pilot projects within a wider programme) can aid robust evaluation.
- Full impact evaluation may not be possible for some complex policy interventions, especially where these are delivered over relatively short timescales. Scoping during the policy design phase what is possible for an evaluation to deliver would be helpful.
- Complex policy interventions often require the involvement of diverse stakeholder groups, which means that different expectations, roles and views on objectives and progress will need to be considered and time needs to be allocated to getting agreement on objectives and evaluation.
- Time is required to develop a good working relationship with the project manager to ensure that any issues around contrasting views on project boards are managed. Time available may be affected by tight project timeframes.

Time is a key element in complexity and must be taken into account in designing evaluation:
- Timescales of delivery (activities and outputs) may differ from intervention outcomes and impacts; many impacts, especially in natural environment initiatives, cannot be detected over time periods of less than 5 years and in some cases decades. Where possible, therefore, longer-term monitoring should build on existing data and plan for the re-assessment of key indicators after the funded intervention has completed.
- An effective evaluation is likely to require an evaluation framework supported by, for example, a clear logic model. Given the potential for delays between activities and outcomes and impacts a theory of change model(s) is a useful approach, accompanied by mechanisms for testing/validating the theory of change.

The appropriate choice of methods is vital in complex evaluations:
- Qualitative and mixed methods are well-suited to addressing complexity in nexus-related evaluations.
- The use of existing national datasets and centralised analysis where possible can help support effective, robust and efficient evaluation at both programme and local levels.
- Self-reported data and locally specific indicators can play a useful role; however, such approaches require support and facilitation, and therefore resources, and may result in inconsistent data.
- Careful consideration is needed in the commissioning and design of bespoke IT systems for short-term policy interventions to ensure that they are proportionate and provide value for money, taking into account the design, implementation, maintenance, and support costs.
- Explicit options appraisal in complex policy development (ex ante assessment) can help inform counterfactual analysis (ex post) providing clear linkage between the different types of assessment/ evaluation.

What kinds of situations bring complexity and where do they occur?

- What we might call 'intrinsic' complexity (because of the subject matter, i.e. issue and impact related) exists equally for nexus-related policy interventions at EU and national/regional/local levels.
- The strength or dynamism of the governance and policy making context, dictates the 'extrinsic' complexity – that exerted not by the subject matter but by the complex web of interrelationship of stakeholders and processes (i.e. policy/response related complexity). In the UK examples, where policy making is more in flux – because of increasing devolved responsibilities to multiple stakeholders – this extrinsic complexity is enhanced. In the EU, it is systematised through rigid processes and frameworks.
- Evaluations of policy interventions where policy is in flux have to deal with very different contexts and enhanced complexity compared to those where there is a strong policy cycle and an evaluation's purpose is not only clear, but explicitly prescribed.

Making evaluation count is challenging but there are factors than can assist

Anticipating the varied challenges in different situations can be helpful:
- High level of instrumental use is seen in EU policy evaluations, because they are designed to deliver that within a strong policy cycle.
- Much of UK environmental policy making exhibits a high degree of flux – more typical of a system stewardship model of policy making/governance where policy, design and implementation are in a state of flux, than a typical policy cycle. Consequently, evaluation has to be more nimble and flexible to respond to ongoing changes in policy purpose, design and implementation.
- Evaluation can have influence in a more indirect way than instrumental – i.e. through conceptual, strategic or process influence and these are more likely in a system stewardship model of policy making than instrumental.
- It is often difficult to know what does happen to evaluations because there is little interaction with the evaluators once they are complete, owing to the contractual nature of the projects reviewed.

Where did this create challenges?

While the evaluations examined were largely fit for purpose there were examples that:
- Lacked a clear policy framework within which to work (i.e. programme level initiatives). This also makes it difficult to understanding how the evaluation was used (if at all) since the policy context is absent.
- Lacked clarity regarding policy and evaluation objectives. From the EU to national policy interventions down to programme level interventions and initiatives this was increasingly the case, reflecting the weak policy cycle (or system stewardship) context for those evaluations.
- Were limited in their use of methods. Because evaluators do what they are asked to do by commissioners of evaluations, there may be limited scope to bring in novel approaches or methods. The methods being used are ones clients are familiar with and understand and that can be used readily for quick evaluations on small budgets.
- Had relatively small budgets, so qualitative methods and theory based approaches were most appropriate, especially where the evaluation may have been the first time an explicit theory of change had been elaborated.

How could these findings be applied in the design of future evaluations?

All of the above considerations have implications for commissioners of evaluations in relation to complex policy interventions around the nexus:

- Commissioners and evaluators need to be aware that an assumption of a traditional policy cycle (however fuzzy that may have been in practice) may no longer be appropriate and that evaluation therefore is less likely to have direct, instrumental use than might have been anticipated.
- Rather than a fuzzy policy cycle, if policy is in a constant state of flux (system stewardship) the purpose, design and implementation of policy are all potentially moving targets, which make it harder to pin down evaluation objectives than when the purpose (objective) of policy is clear.
- This meta-evaluation provides substantive evidence of this type of policy flux and the challenges the evaluations in those situations faced along with the need to tailor make evaluations each time to those circumstances.
- For such evaluations to have impact increasingly evaluators will need to be nimble and responsive to changing policy purpose, design and implementation and understand where within this new system stewardship 'policy triangle' evaluation could impact most effectively.

Key questions for new evaluations

- What is the nature of the policy context in which your evaluation is being carried out? Would you describe it as evolving, stable, unclear, high profile?
- How far are the objectives of the policy or intervention/initiative clear and amenable to evaluation? Are the expected outcomes and impacts clear?
- How far are the objectives of the evaluation clear and achievable given the nature/timing of the policy/intervention/initiative and the resources of the evaluation?
- Are there multiple stakeholders involved as part of the steering group for the policy intervention/initiative? How far is there consensus across perspectives? Are there clear mechanisms in place to enable management of different perspectives?
- Is there a clear and active Project Manager for the evaluation?
- What are the expectations of the client in relation to the ability of the evaluation to evaluate longer term impacts?
- What types of complexity are most relevant to the evaluation (with reference to the four categories and sub-categories in first section)?
- To what extent do you think your methods are appropriate for evaluating these complexities? What strategies can you use to address these specific aspects of complexity?
- What types of impact are expected by your evaluation? How will the client assess whether they have been realised?
- How can you improve the impact of your evaluation? Where are the points of influence within the evaluation?

References and further information

Collingwood Environmental Planning Ltd has, over the last 10 years (2006-2016), undertaken an extensive range of evaluations in the natural environment arena, e.g. for Defra, Environment Agency, Natural England, Research Councils, Scottish Government, European Commission, Natural Resources Wales, OECD, etc. This EPPN is based on that work.

Hallsworth, M. (2011) System Stewardship - The future of policy making? London: Institute for Government. http://observgo.uquebec.ca/observgo/fichiers/85195_aeepp1.pdf

Sheate, W.R.,Twigger-Ross, C., Papadopoulou, L., Sadauskis, R., White, O., Orr, P., Phillips, P., and Eales, R. (2016) Learning lessons for evaluating complexity across the nexus: a meta-evaluation of CEP projects, Final Report to CECAN, November 2016. Available at http://cecan.ac.uk/news/learning-lessons-from-past-interventions-collingwood-environmental-planning

www.cecan.ac.uk / cecan@surrey.ac.uk / +44 (0) 1483 682769

The Centre for the Evaluation of Complexity Across the Nexus (CECAN) is a £3m national research centre hosted by the University of Surrey, bringing together experts to address some of the greatest issues in policy making and evaluation.
This Evaluation Policy and Practice Note was written by Dr Bill Sheate and Dr Clare Twigger-Ross, Collingwood Environmental Planning Ltd. Contact: w.sheate@cep.co.uk and c.twigger-ross@cep.co.uk