



## Evaluating a large energy innovation portfolio

*Nico Jabin, Robert Rutherford, Dept for Business Energy and Industrial Strategy, UK  
(nico.jabin@beis.gov.uk)*

### EXTENDED ABSTRACT

In 2016 the UK government launched the Energy Innovation Programme, a five-year, £505 million programme that aims to accelerate the commercialisation of innovative clean-energy technologies and processes into the 2020s and 2030s. This abstract discusses the approach the Department for Business, Energy and Industrial Strategy (BEIS) has taken to evaluating this portfolio, and the lessons it has learned during the process.

### Introduction / Background

The Energy Innovation Portfolio (EIP) is managed by the Department for Business, Energy and Industrial Strategy (BEIS), includes over 70 programmes within 6 technology themes and has funded around 450 projects. The themes cover energy-related sectors, from industrial decarbonisation to smart systems, from nuclear energy to renewables. Most programmes support technologies across the so-called ‘valley of death’ in the innovation lifecycle, but some take different approaches, e.g., funding a research centre. Programmes differ widely in various ways, e.g., scale of projects and funding, supported organisations, stage of innovation. BEIS employs over 100 engineers, project managers and policy leads to manage programmes. The notion of evaluation as understood by government social researchers<sup>1</sup> is not well ingrained.

The challenges in evaluating this portfolio arise from its size, complexity and often formative nature, from the long innovation timescales coupled with the desire for early evidence, the need for risk and failure, the complex and dynamic systems in which innovation takes place<sup>2</sup>, and the singularity of many innovations.

### Methodology

The initial phase of development of the EIP evaluation aimed to resolve a dilemma: responding to criticism levelled by the NAO against the lack of evaluation of the EIP’s predecessor, and the difficulty of setting up an evaluation team that had little to evaluate in the programmes’ early stages. As a result, the evaluation function grew incrementally, moving from core monitoring to evaluations of priority programmes, to developing a learning function and, finally, to engaging with the systemic nature of innovation.

The first step, in the second year of the portfolio, was the development and embedding of portfolio-wide key performance indicators (KPIs). The following year, theme and programme leads jointly prioritised, and evaluators would lead evaluations of the most important (Level 1) programmes, for Level 2 programmes the programme lead was to manage, and Level 3 programmes were monitored using the portfolio KPIs.

In the EIP’s fourth year, planning started on the implementation of a suite of commissioned evaluation projects, which were scoped and designed with the assistance of a technical support contract. Because most

<sup>1</sup> HM Treasury. Magenta Book: Central Government Guidance on Evaluation. 2020.

<sup>2</sup> See for example: Geels, Frank W. “Socio-Technical Transitions to Sustainability: A Review of Criticisms and Elaborations of the Multi-Level Perspective.” *Current Opinion in Environmental Sustainability*, vol. 39, 2019, pp. 187–201. Perrin, Burt. “How to — and How Not to — Evaluate Innovation.” *Evaluation*, vol. 8, no. 1, 2002, pp. 13–28.

programme evaluations were still in their infancy or in planning, the evaluation team also developed a ‘learning’ work strand. This aimed to identify, using a light-touch approach, policy-relevant findings from innovation projects within a specific theme and communicate these to internal stakeholders.

In 2020, finally, the KPI data collection was centralised and the portfolio-level approach to synthesis, as well as the evaluation approach for the EIP’s successor, the Net Zero Innovation Portfolio, developed.

## Results

The EIP evaluation approach, now, focuses largely on the evaluations of Level 1 and 2 programmes, accompanied by a portfolio-level KPI approach. Most Level 1 evaluations are at an early, or at scoping or commissioning stage, partly due to late programme starts and partly due to limits on resources. Our support to Level 2 evaluations is greater than we anticipated, and we have not yet rolled out as much support as is necessary to make this a successful working model. Level 3 evaluations, unsurprisingly, take up little time.

The evaluations address typical areas of concern: improving processes throughout the innovation support chain; and the outputs, outcomes and impacts the innovation work has achieved. Most evaluations use a theory-based design to assess impact<sup>3</sup>, only very few programmes lend themselves to meaningful counterfactual approaches. Evaluations usually assess impact at 3-year follow-up.

Where evaluations have produced outputs already (usually interim reports), they have been very well received by delivery and linked policy teams, and this has greatly increased buy-in to evaluation work.

KPI data collection was originally the responsibility of programme leads, but following concerns about quality and regularity, we have last year moved to a centrally run survey. This has achieved a 95% response rate, and reasonable data quality. It will feed directly into the developing portfolio synthesis, which will bring together programme evaluation findings and KPIs. The KPI data has enabled a comprehensive view of the portfolio, which has improved communication and insights, but we have found it difficult to answer the most interesting questions on impact through these indicators.

The ‘learning’ project has provided some valuable insights directly to policy teams, and has definitely received support for its interim communication, but it has been difficult to resource.

## Conclusion & discussion

Our reflections can be divided into four groups. First, it is important to plan resources carefully from the beginning. While innovation programmes may take time to start, so does evaluation. Developing and embedding processes, prioritising and planning, developing a portfolio-level approach, scoping at programme level all take time, and pushing these elements into the future is storing up difficulties. At the same time, value for money from evaluation can be improved from combining related programmes for evaluation purposes, and using portfolio-level approaches to answer cross-cutting questions and enable counterfactual approaches.

Second, embedding evaluation is difficult. It requires not only sufficiently resourced programme leads, but also ongoing senior support, processes, evaluation support and engagement. This is true for larger ambitions such as programme lead-led evaluations and for easier exercises such as KPI data collection. Embedding evaluation is best envisaged as cultural transformation project.

Third: Evaluation users do not want to wait until 3-years post-programme. Plan for your ability to present interim findings, case studies and summary updates to senior staff for key timepoints such as spending reviews or preparation for subsequent funding rounds, and at regular intervals.

Finally, there are opportunities for even greater impact in the future: building our evaluations on a better theoretical foundation of expected innovation impacts; aiming to contribute directly to the academic literature; and reviewing what data we could make available to academic researchers.

---

<sup>3</sup> E.g., Barr, Julian, et al. [GCRF Foundation Stage Evaluation: Final Report](#). Department for Business, Energy and Industrial Strategy, Nov. 2018.