
SRA Conference: Using Innovative Theory Based Evaluation Methods in Energy Policy

Emily Sumner, Energy Evaluation, emily.sumner@beis.gov.uk

Adam Stiles, Energy & Security, adam.stiles@beis.gov.uk


Department for
Business, Energy
& Industrial Strategy

Overview

Aim:

To give you an insight into how BEIS evaluate some of their energy policies and illustrating this using some case studies.

This presentation will cover the following:

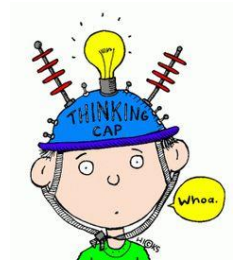
- Evaluation at BEIS and the energy evaluation team
- How do we choose an appropriate evaluation method?
- What do we mean by theory based evaluation methods? (Including case study examples).
- Practicalities of using innovative methods and communicating findings

Evaluation at BEIS

Evaluation at BEIS

Evaluation at BEIS is important, as it provides **learning** and **accountability** of spend, and allows us to:

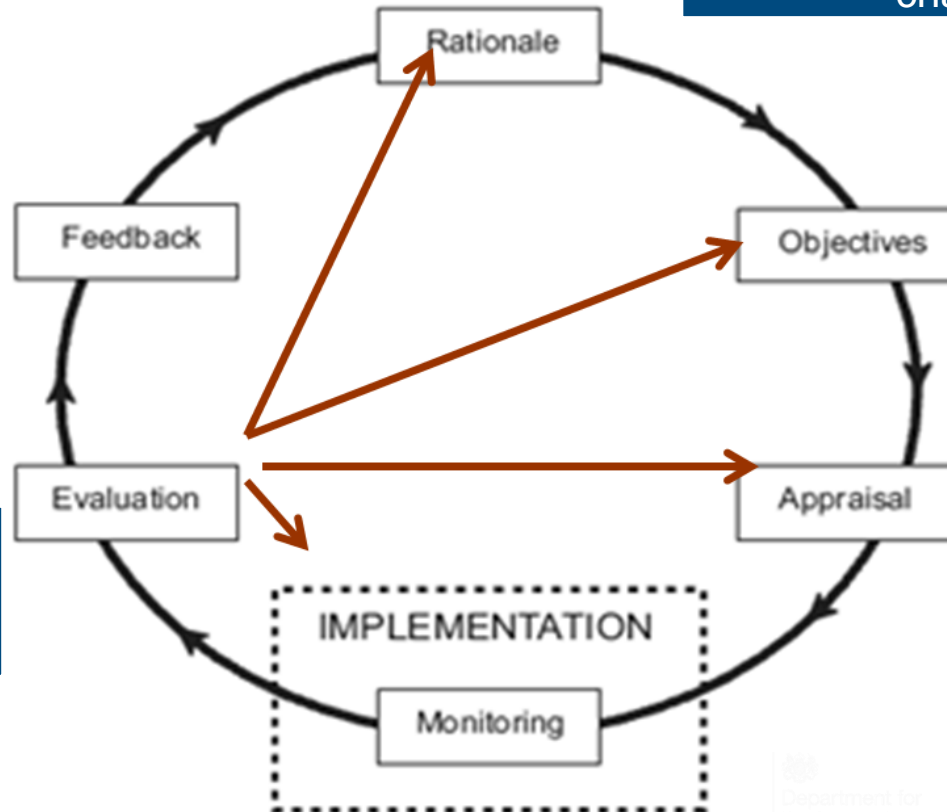
- Improve policies through **learning** and **understanding**, and **managing risks**, especially where there is **uncertainty**.
- **Transparency** of government spend, especially for high spend policies.
- **Build knowledge** and collective understanding, for example where a pilot scheme might be rolled out.
- Inform and **improve policy decisions**.



Where does evaluation feed in?

Learning is reflected in future decisions

Planned and implemented early, (e.g. clear logic model, theory of change)



In sync with policy timelines

Department for
Business, Energy
& Industrial Strategy

What energy-related areas are we evaluating?

- Smart Metres
- Heat Networks
- Renewable energy support
- Flexibility electricity capacity
- Electricity demand
- Energy Innovation
- Business and Non-Domestic energy
- Energy efficiency
- Climate change
- Renewable Heat

How do we choose an appropriate evaluation method?

Different approaches to evaluation

Experimental – Allocation to treatment and control groups, providing a comparison group to demonstrate an effect.

Quasi-experimental – Allocation isn't possible, constructing a comparison group using propensity score matching, difference in differences.

Theoretical – Where the above isn't possible / or a more detailed understanding is required. How and why policy causes change.

These methods aren't mutually exclusive....

Why are theory based methods appropriate?

Lack of a control grp

- Many policies lack a natural **control group** or counterfactual as policies are rolled out across England/GB/UK
- Or sample sizes are too small to form a robust counterfactual.

Applicability

- TB methods provides **a robust alternative** to not being able to apply experimental approaches, and prevents misapplication of experimental approaches. Provides robust and rigorous evaluation to inform policy.

New, complex policy areas

- Many energy policies are **new, innovative, and complex**, with lots of evidence gaps. TBM's allow for a greater focus on context and understanding what's working for who and under what circumstances.

Depth of insight

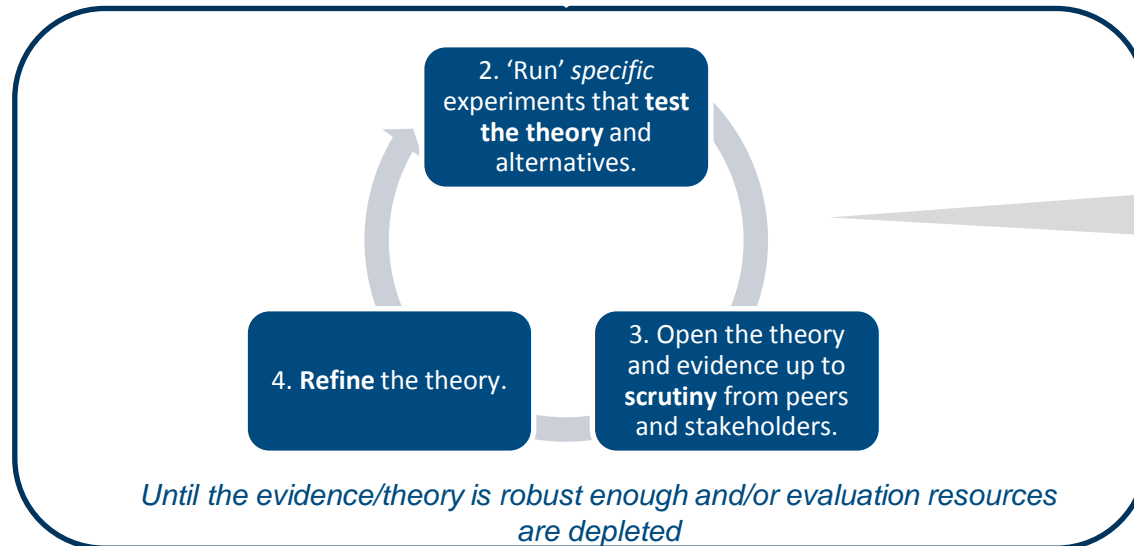
- Theory based approaches allow us to provide a **more in-depth, theoretical approach** to understand to start to build the evidence base.

What do we mean by theory based methods

What do we mean by theory based evaluation?

1. Develop a theory of the policy/scheme
How will this intervention lead to the outcomes required?
What are the alternative theories of how the outcome could come about?

Develop theory:
E.g. a linear theory of change, or a more nuanced theory (e.g. for different sub-groups) .



Test & refine theory:
Test initial theory using a range of evidence, could be primary research, or using existing data.

5. Consolidate the new understanding of how the policy works/does not work. Make an assessment of what the results suggests about the impact of the scheme.

New theory is used for future decision making.

Case Study Examples of Theory Based Methods


Department for
Business, Energy
& Industrial Strategy

Electricity demand reduction pilot evaluation

EDR Pilot - Provides incentives to organisations to implement energy efficiency measures in their buildings, to reduce electricity use at peak times.

Want to use theory based methods to assess the impact of the scheme:

[Q] Is the policy leading to greater reduction in electricity at peak times, compared to if there was no government incentive?

Why experimental methods were not deemed appropriate:

- No random allocation to treatment group
- No meaningful counterfactual
- Sample size too small to allow statistical analysis (but allowing census).

How was this done?

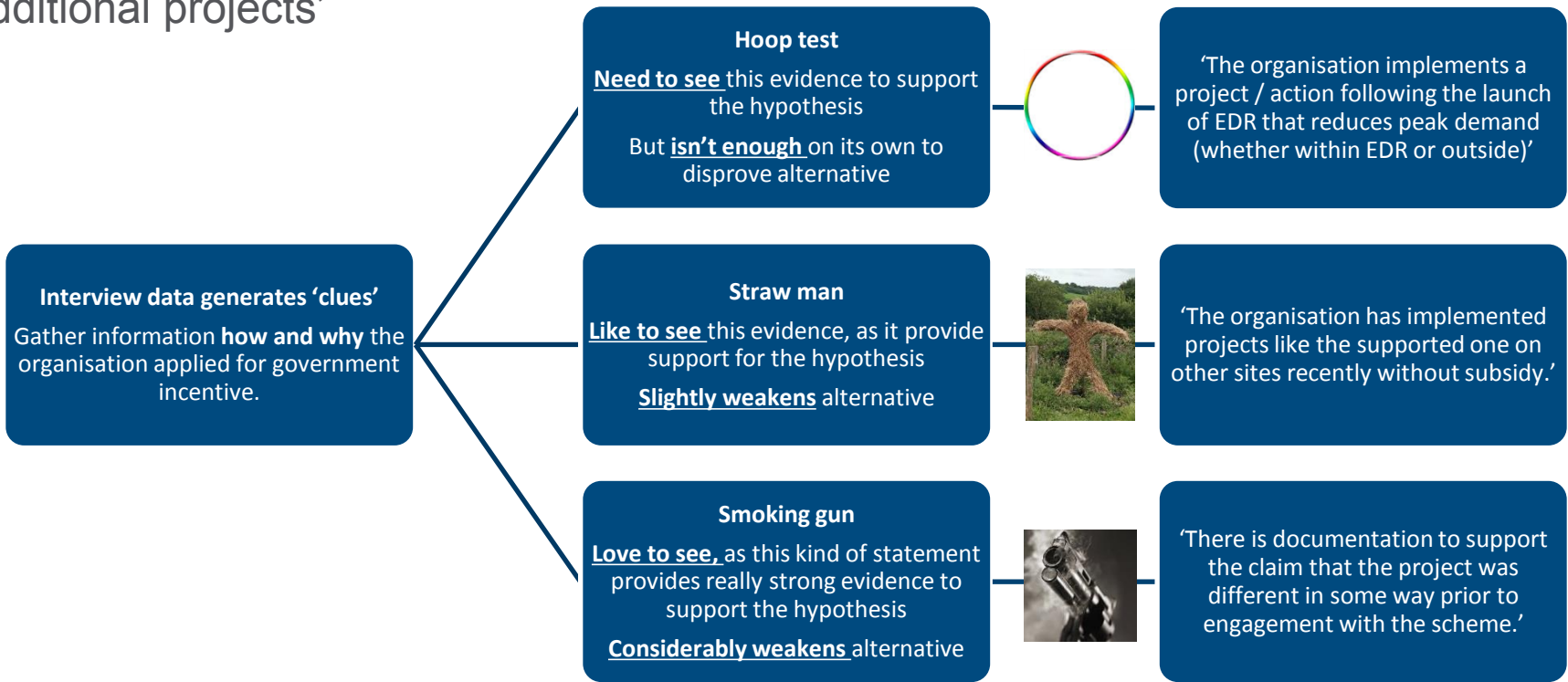
- Range of interviews - different levels of engagement at different stages of the pilot process, (triangulated with case notes).
- Process tracing to gain evidence of the level of additionality


Department for
Business, Energy
& Industrial Strategy

Electricity demand reduction pilot evaluation

Case based approach to causal inference, that uses ‘clues’ to support and/or oppose the hypothesis: ‘the incentive led to additional projects’

Doubly decisive:
Confirms hypothesis and eliminates the alternative.
This wasn't used for EDR.

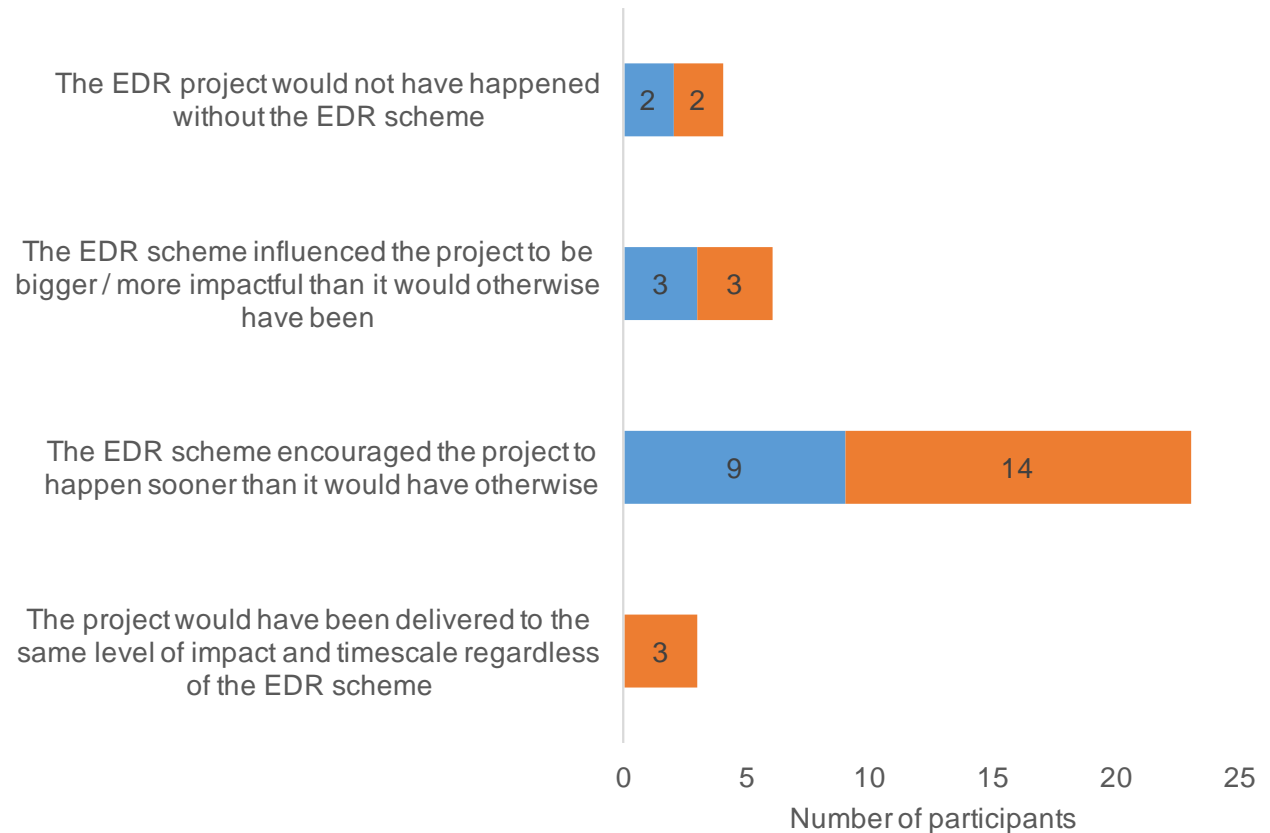


Electricity demand reduction pilot evaluation

Four classifications of 'additionality'

Majority of projects were additional on some level, with 3 claiming no additionality (projects would have been implemented regardless).

External projects – on the whole, the scheme was effective in deterring non-additional projects.



Note: some had multiple projects in the scheme, which were rated differently in terms of attribution.

■ Phase I ■ Phase II

Transitional arrangements for DSR evaluation

Similar problem to EDR evaluation

no counterfactual

small sample size

no eligibility criteria

lobbying bias

qualitative outcomes

other revenue streams

The solution?

Theory based evaluation based on contribution not attribution

Want to assess the extent to which the TA contributes to desired outcomes, but also if other explanations fit better for each participant

Department for
Business, Energy
& Industrial Strategy

Policy Objective: DSR Capacity made available in the short term

Policy Contribution Hypothesis:

But some followed this contribution story, demonstrating that the policy did contribute to this objective i.e. additional

Alternative Hypothesis:

Most capacity in the policy followed this contribution story i.e. was non-additional

Business, Energy
& Industrial Strategy

Practicalities of using innovative methods


Department for
Business, Energy
& Industrial Strategy

Practical Considerations - Communicating

Working as a government social researcher, you can sometimes face challenges when **communicating findings**. This becomes even more difficult when using complex methods to policy officials who:

- Often don't have analytical or evaluation backgrounds
- Aren't familiar with the terminology
- Are resource constrained, so have limited time (e.g. engaging in detailed findings, and investing time in theory development).
- Can have a higher reliance on numbers/ find numbers easier to interpret

Some key tips:

Being creative - using diagrams, presentations/ workshops, flow charts.

Repeating key messages

Keeping policy involved throughout the process, not just when presenting findings (e.g. steering groups).

Not going into methodological detail, focus on key implications for policy.
Digesting findings into manageable chunks.

Using simple terminology, or explaining in simple language (e.g. theory of change).

Other Practical Considerations

Having to **upskill ourselves** on new and emerging methods, that aren't necessarily well established, e.g. realist methods, flexible approach.

Lack of established **quality assurance** of theory based or mixed method approaches, whereas the Maryland is recognised and easy to communicate.

Some methods can be **demanding of policy's time** – for example defining and refining theory of change, requiring them to be actively involved.

A strength of the approach is that rich detailed findings can **influence later stages in the evaluation**; these decision points need to be built in.

Some policies have **very long term goals** (e.g. climate change targets), so it can be hard to test the impact. So often focus on the outputs and outcomes.