



Cabinet Office

Theory-Based Designs

Evaluation Task Force Academy

Hello



ETF Evaluation Academy

Module 1: Introduction to Evaluation

Module 2: Developing a Theory of Change

Module 3: Scoping an Evaluation

Module 4: Process Evaluation

Module 5: Impact Evaluation - Experimental Designs

Module 6: Impact Evaluation - Quasi-Experimental Designs

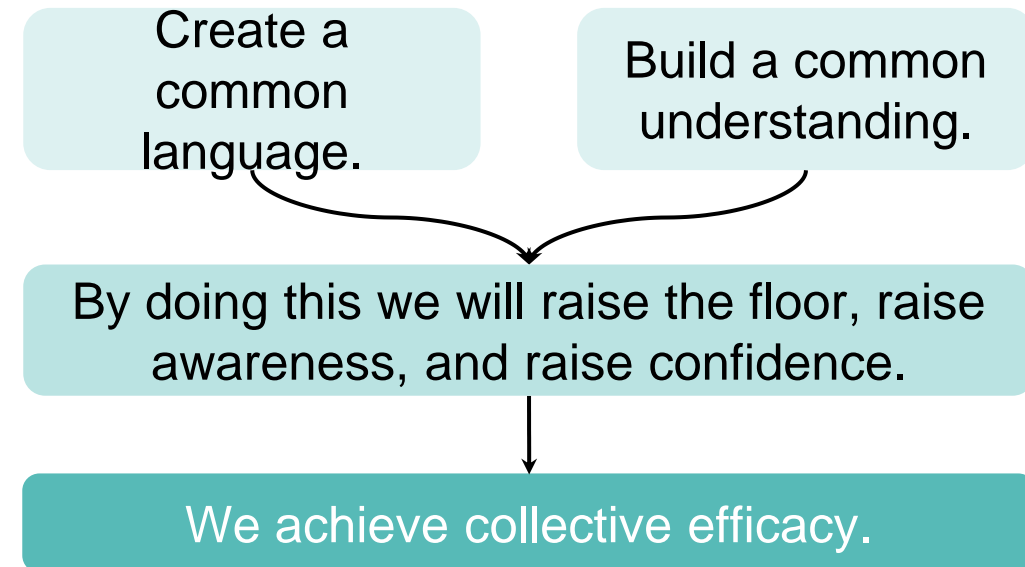
Module 7: Impact Evaluation - Theory-Based Designs

Module 8: Value for Money Evaluation

Module 9: Planning and Managing an Evaluation

Module 10: Communicating Evidence and Decision Making










The Evaluation Academy will upskill analysts across HMG departments in key evaluation methodologies and evaluation management techniques and will result in better and more evaluation across HMG.



Module 7: Overview of contents

- [Learning outcomes](#)
- [Hierarchy of evidence](#)
- [What are theory-based designs?](#)
- [How do you run a theory-based evaluation?](#)
- [How do I run a Realist Analysis and Contribution Analysis?](#)
- [Additional TBE methods](#)
- [Benefits, risks and uses of TBEs](#)
- [Advocacy and application of learning](#)

Learning outcomes

-  I can **explain** the role and value of key theory-based impact evaluation designs (realist analysis, contribution analysis)
-  I can **describe** other theory-based designs (process tracing and qualitative comparative analysis)
-  I can **identify** the types of evaluation questions that TBE designs can and cannot answer
-  I can **explain** how TBE designs build on a ToC and can work alongside other evaluation methods
-  I can **explain** the key things that need to be considered when preparing and running each TBE design
-  I can **explain** the benefits and risks of different TBE designs
-  I can **contrast** the most appropriate TBE method(s) to use based on the policy context
-  I can **critically assess** the findings of a TBE
-  I can **advocate** for including TBE's across the policy cycle

ETF Evaluation Academy: Hierarchy of evidence

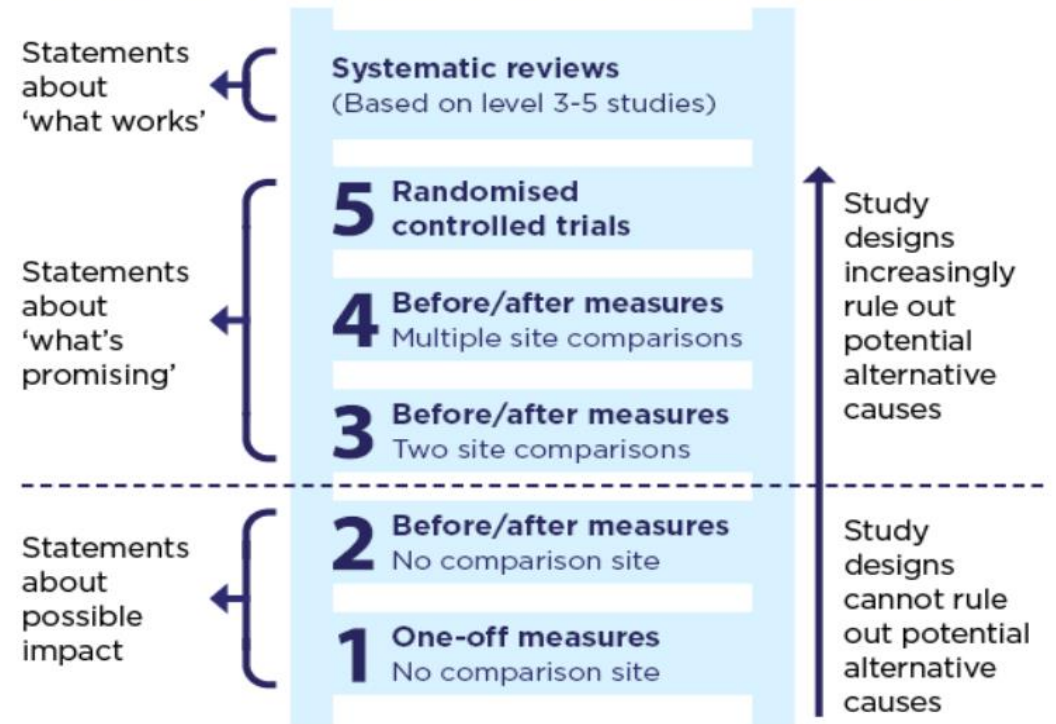
The Nesta Standards of Evidence

The objective of developing Standards of Evidence is to help us know how confident we can be in the evidence provided to show that an intervention is having a positive impact.



Ladder of evidence

How can we be confident our activity makes a difference?



What are theory-based designs?

Have you ever thought:

You may already be applying some of the principles of TBEs in your day to day work.



How much do I have confidence the results of my impact evaluation?



What else could have caused a change I observed?



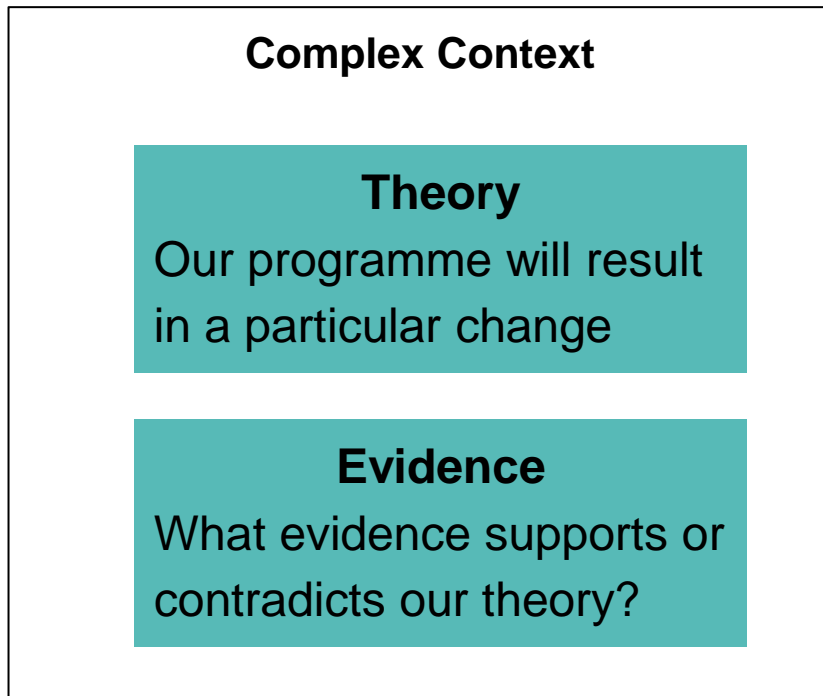
What information would I need to see to be more confident in the results?



Can I apply this programme to a different group of people and see the same result?

What are TBEs and when would I run one?

TBEs are an umbrella set of techniques that can be used as standalone evaluation tools or to strengthen the understanding you gain through other evaluation methods.



Awareness campaign



General increase in digital literacy



Civil society support



Rates of digital filing

TBEs are suited to complex domains



Ensure UK contributions to European Space Agency are properly evaluated



Use Theory of Change to develop evaluation framework



Validate pathways to impact

- ✓ Complex intervention
- ✓ Long timescales
- ✓ No counterfactual available

What kind of questions can a TBE answer?

Imagine you are running a programme to improve attendance in schools. You ran it in Bristol and saw positive results. You ran the same programme in Middlesbrough and saw no positive impact.

Did our programme *contribute* to improvements in school attendance in Bristol?

Did the programme work *differently* for different groups of students in Middlesbrough?

Why did our school attendance programme work in Bristol, but not in Middlesbrough?

Where else might it work and how should our attendance programme be adapted to scale it to new cities?

What would need to be true for this to improve attendance in *Bristol* in the future?

How did my programme improve attendance in Bristol?

What was the degree or magnitude of the change it caused?

TBEs in HMG



Cabinet Office: GREAT Campaign

Does the campaign generate returns for the UK economy?

Does the campaign help the country present itself positively to existing and future target markets?



DLUHC: Supported Housing

Are existing powers sufficient to address poor quality?

Which interventions (from five LAs) had the greatest impact?



MoJ: Reforms to HMCT

Has reform delivered its intended objectives (a just, proportionate, and accessible system)?

What are the specific lessons that can be learnt from reform?



Scottish Gov: Inward investment

Does investment promotion support by enterprise agencies support job creation and spillover benefits in the Scottish economy?



Activity: In breakout groups discuss a programme you've worked on that might be suitable for a TBE.

Common Theory-Based Impact Evaluation methods

In this module, we'll mostly focus on Realist Evaluation and Contribution Analysis.

Method	Description
Realist Evaluation	Specific, hypothesised causal 'mechanisms' for an 'outcome' are articulated in 'context' and evidence gathered for each.
Contribution Analysis	Step-by-step process used to examine if an intervention has contributed to an observed outcome by exploring a range of evidence for the Theory of Change.
Qualitative Comparative Analysis	Used to compare multiple cases and systematically understand patterns of characteristics associated with desired or undesired outcomes based on qualitative knowledge.
Process Tracing	A structured method examining a single case of change to test whether a hypothesised causal mechanism, such as that proposed by the Theory of Change, explains the outcome.
Bayesian Updating	Probabilities of a small number of contribution claims are estimated prior to observation then tested.
Contribution Tracing	Participatory mixed-method to establish the validity of contribution claims with explicit criteria to guide evaluators in data collection and Bayesian updating to quantify the level of confidence in a claim.
Outcome Harvesting	Collects evidence of change and then works back to assess contribution to the change.
Most-significant Change	Involves collection of significant change stories from the field and systematic selection of the most significant by panels of stakeholders.

TBEs compared to Process Evaluations

PE

How a programme is implemented and run, how that implementation affects outcomes and what good implementation looks like.

Did the students in Bristol *perceive* an impact?

How was the programme *implemented*? Were there deviations from the original plan?

Did the programme reach the target group?

Ready for impact evaluation? What changes are needed to do this?

BOTH

Use quant and qual methods to provide depth and richness.

Why did our school attendance programme work in Bristol, but not in Middlesbrough?

Where else might it work and how should our attendance programme be adapted to scale it to new cities?

TBE

Why and how the programme works or *contributes* to an outcome.

Did our programme *contribute* to improved school attendance in Bristol?

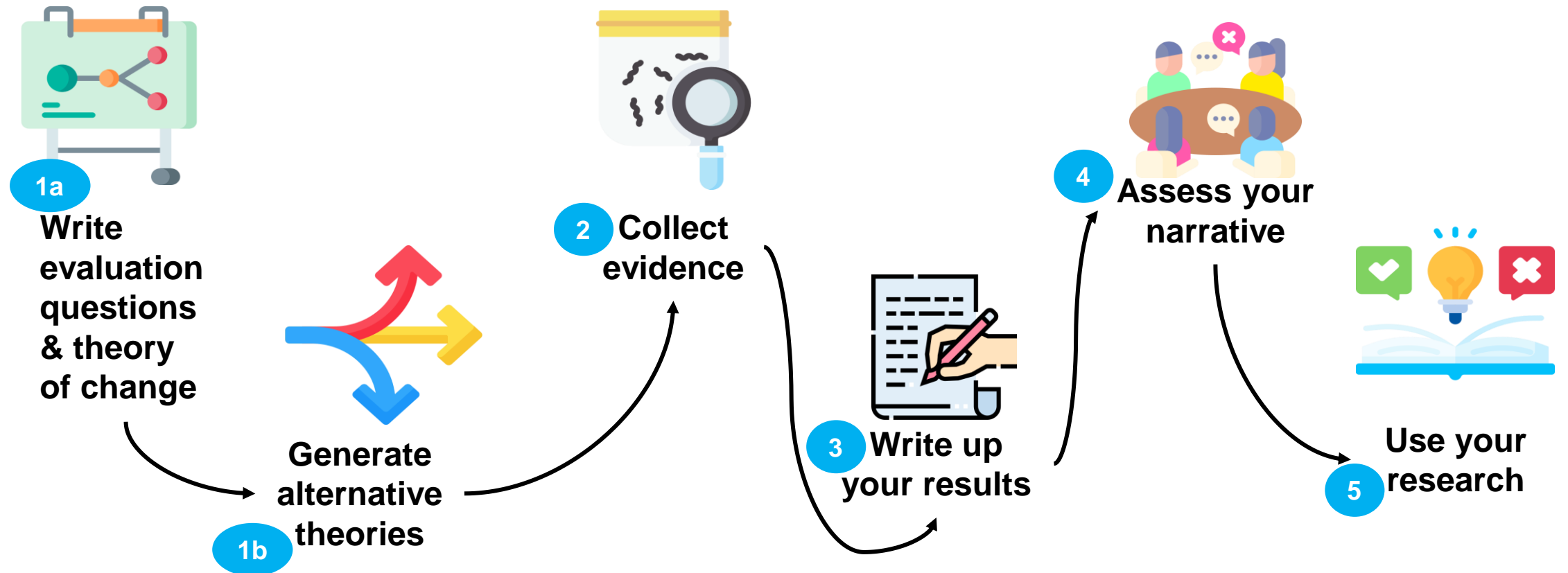
Did the programme work *differently* for different groups of students in Middlesbrough?

How did my programme improve attendance in Bristol?

What would need to be true for this to improve attendance in *Bristol* in the future?

How do you run a Theory Based Evaluation?

How do you run a Theory Based Evaluation?



Case study and evaluation questions

Let's imagine we wanted to use a TBE to evaluate this ETF Evaluation Academy...



The Evaluation Academy includes 10 modules that teach departmental analysts the required skills to become an effective evaluator.



Desired outcome: an increase in the level and quality of evaluation across HMG



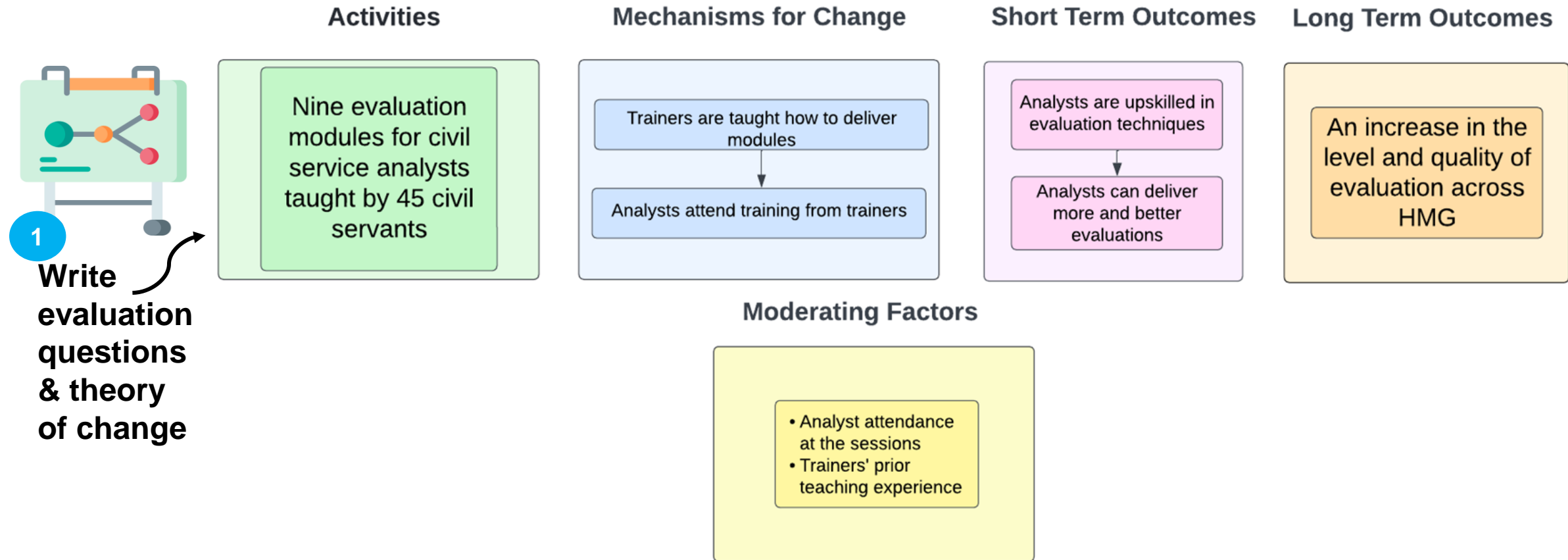
Activity: What are 1-2 possible evaluation questions?

Did the Evaluation Academy *contribute* to increase in the use and quality of evaluation in HMG?

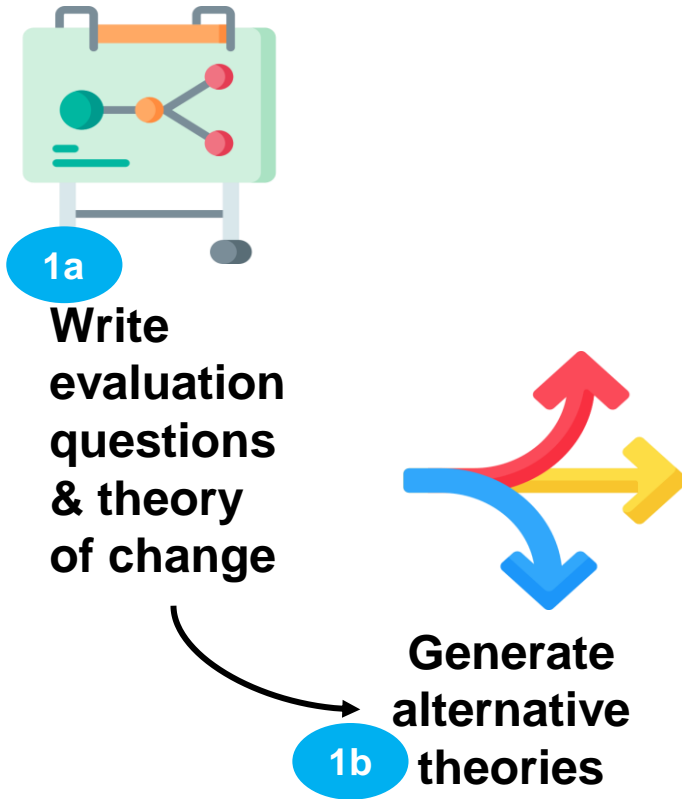
How should the Academy be adapted to scale it to new departments or professions within the Civil Service?

TBEs build off a Theory of Change (ToC)

Start with a ToC and work backwards to identify evidence that supports your theory.



What *alternative theories* could explain the outcome?



The Evaluation Academy is 10 modules that teach departmental analysts the required skills to become an effective evaluator.

Desired outcome: increased use of evaluation methodologies and evaluation management techniques by HMG.



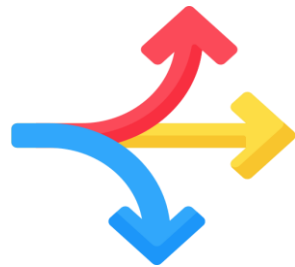
Activity: What alternative theories could explain an increase in the level and quality of evaluation across HMG?

Determine your alternate theories



1a

**Write
evaluation
questions
& theory
of change**



**Generate
alternative
theories**

1b

Theory 1: Participation led to:

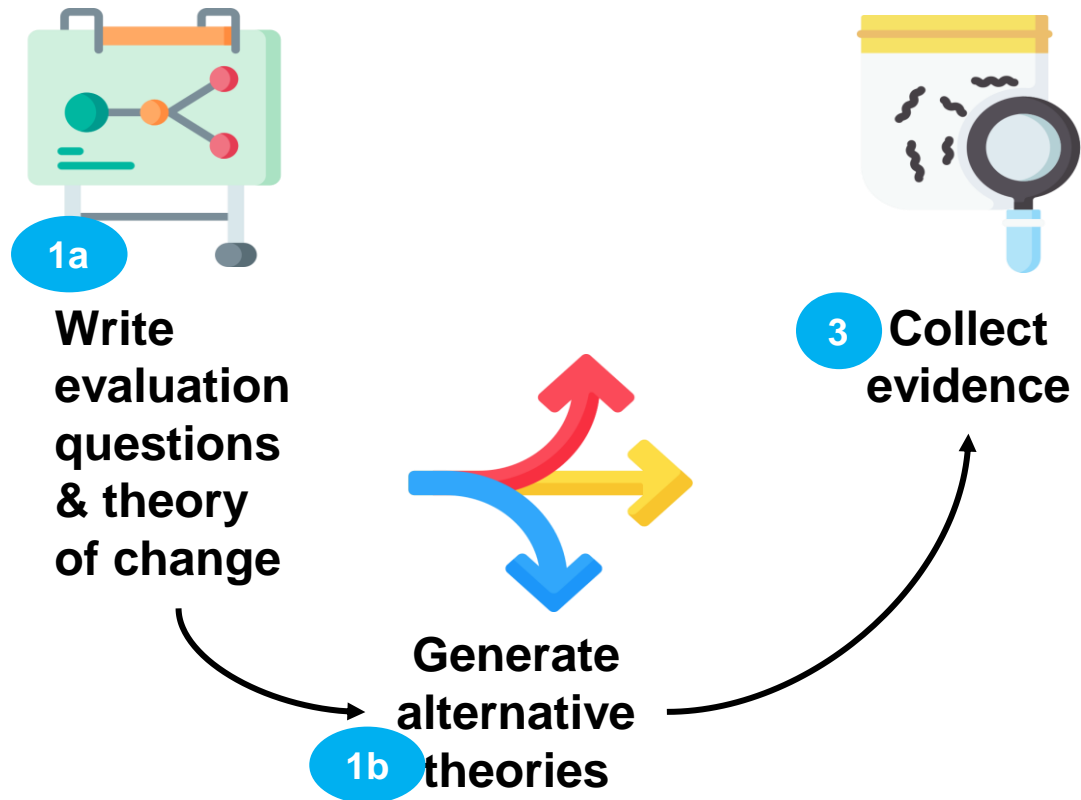
- increased awareness of evaluation methods
- increased ability to advocate
- network of peers to lean on to tackle difficult problems when using evaluation

Theory 2: The ETF's work persuades Senior Government officials of the need for more and better evaluation, and they look for it in all business cases they receive.

Theory 3: New rules in the Spending Review process mandate the inclusion of evaluations of most government programmes.

Theory 4: Government starts recruiting more analysts with an evaluation skillset, leading to more capacity to deliver evaluations.

Gather evidence to investigate your claim



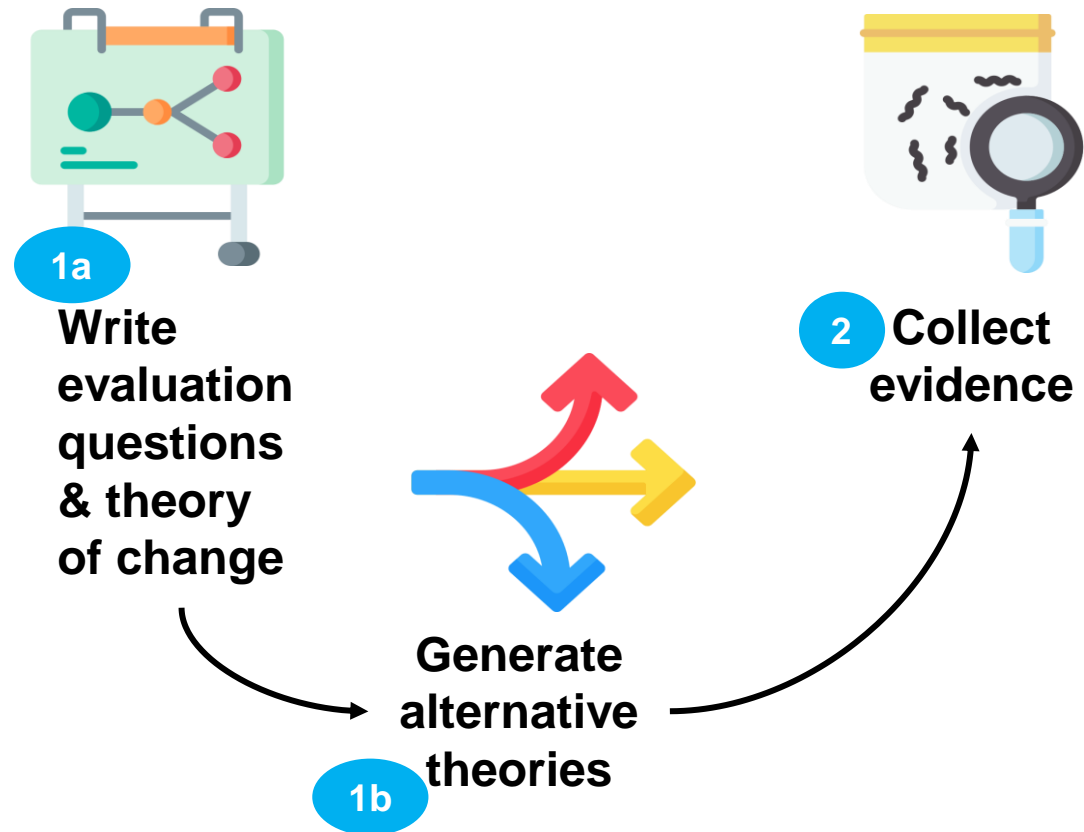
You can use quantitative and qualitative data...

... but remember, don't only look for **confirmatory** evidence!



Activity: What evidence might help us to answer our evaluation questions for the ETF Evaluation Academy?

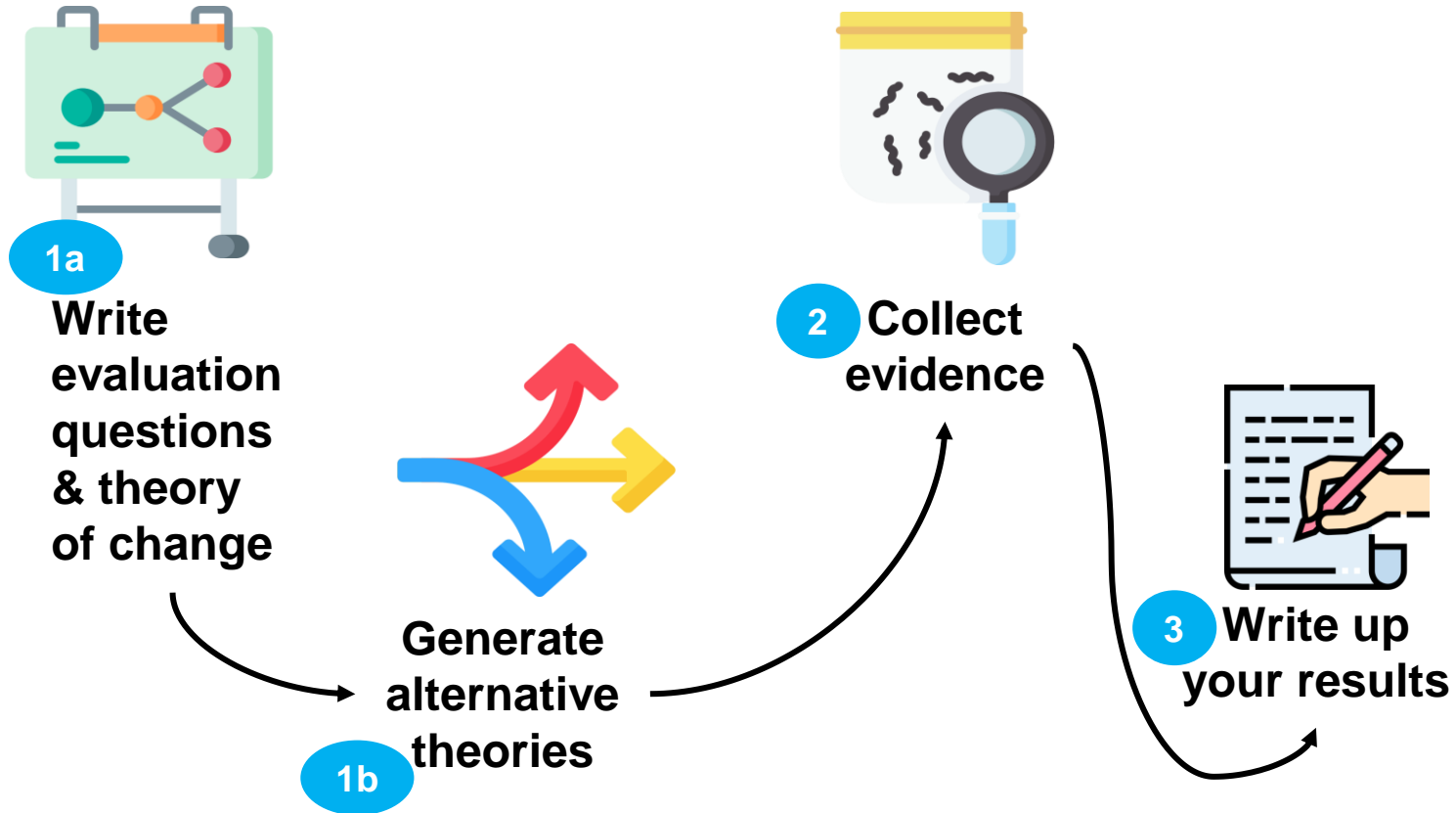
What evidence might help *investigate* the Evaluation Academy's role



Possible evidence

- Interview or survey participants - what are their perceptions of how and if the training contributed to the outcome?
- Create a training delivery timeline and map against the number of evaluations completed.
- Map timeline against the Government spend on evaluation.

Write up your results



Write up your results



3 Write up your results

The format will vary depending on the type of TBE:

Contribution Analysis

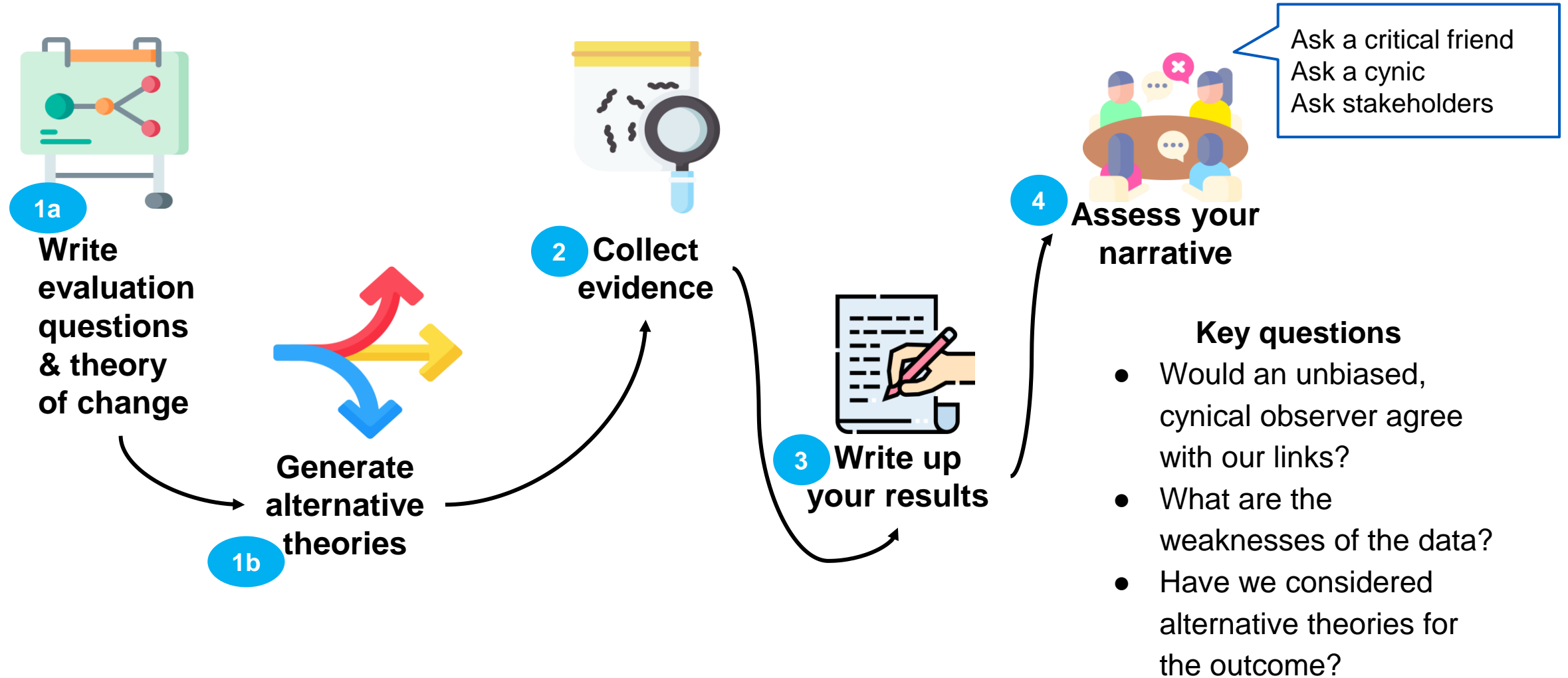
Contribution narrative - written and/or visual depiction of how the programme contributed to a change.

Realist Analysis

CIMO statements - weave together the Context, Intervention, Mechanism and Outcomes.

Civil Servants in the UK government (C) attended 9 training modules connected to evaluation (I), which deepened their expertise and advocacy (M), leading to increased use of evaluation in public policy (O) .

Assess your results



Assess your results



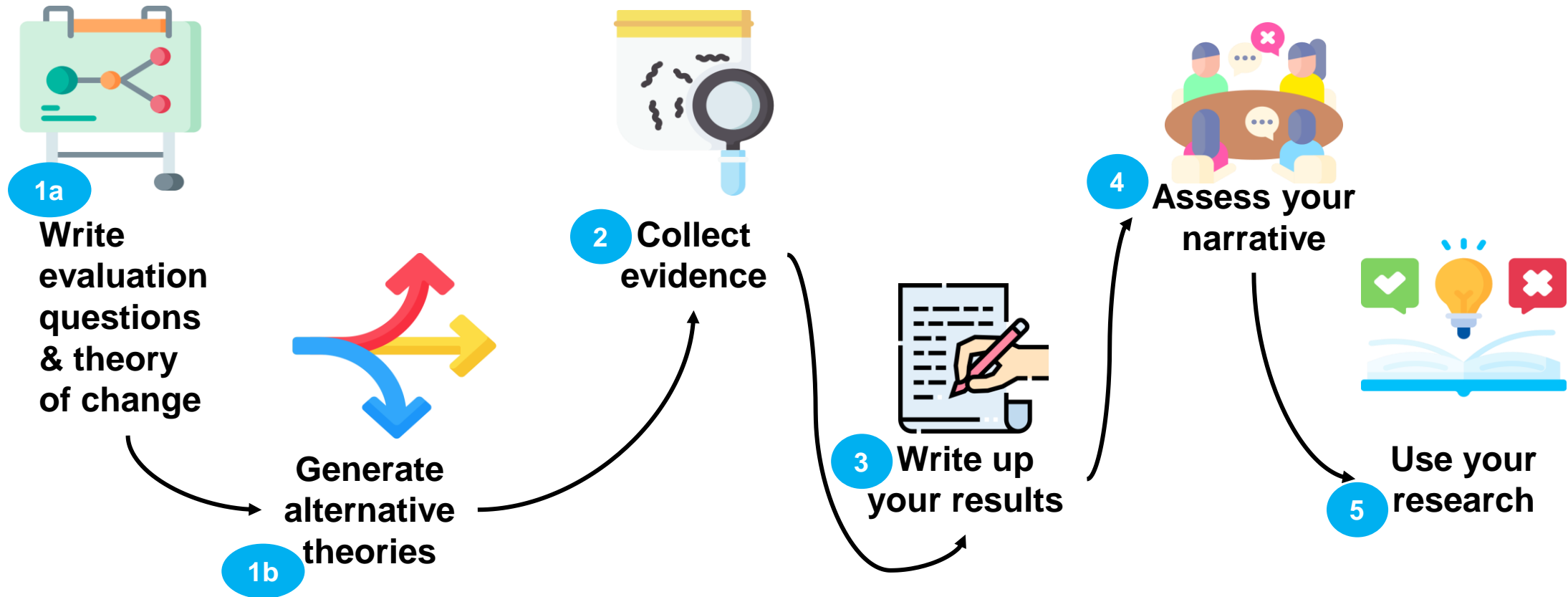
Activity: Who would you invite to be your critical friend and in what forum would you gather their feedback?

Academy participants not included in the TBE.

External experts on evaluation and systems change.

Civil servants who have tried to change other internal systems.

Use your research



Use your research

You are confident the outcome can be linked to your programme

Use your evidence for future policy work.

Strong evidence for components of your programme, weak or no evidence for others

Recommend directing more resources to the most impactful components. Remove or gather evidence for less impactful components.

Uncertainties remain about the model and you can address them

Make a plan to gather further evidence, then revise and strengthen your narrative based on the new evidence.

Uncertainties remain about the model but you can't address them

Stop your programme, or acknowledge it is not research-backed. Consider data collection methods that would evaluate causality.

How do I run a Realist Analysis and a Contribution Analysis?

Primary methods of theory-based evaluation

Realist Analysis

Case Study

Climate Public Private Partnership (CP3), a joint initiative from DfID and BEIS.

An investment fund of £130m to increase low-carbon climate investment in developing countries.

Contribution Analysis

Case study

Coronavirus Community Support Fund (CCSF), funded by DCMS.

The CCSF was targeted at small and medium sized community organisations delivering activities and support to people affected by the COVID-19 crisis.

Realist Analysis

Fundamentals of Realist Analysis (RA)

A way of thinking about your entire evaluation programme.

What works, for whom, in which circumstances, and why?

‘Realist Evaluation is based on realism (a philosophical perspective in which the social world is viewed as real). Therefore, non-observable entities and processes such as culture, class and economic systems can have a real effect on whether programmes work.’

[Public Health England](#)

When would I use an RA?

RA is appropriate when evaluating..

- ✓ Complex programmes with wider learning potential
- ✓ Programmes that sit within a system
- ✓ Programmes that produce mixed outcomes

RA is not appropriate when...

- × The how, why and where programmes work is already understood
- × You are chiefly interested in the effect of the programme

Realist Analysis: Climate Public Private Partnership (CP3)

The Climate Public Private Partnership (CP3) is a joint initiative from DfID and BEIS.

The objective of CP3 is to **increase low-carbon climate investment** in developing countries. These investments are expected to provide commercial returns to the UK Government. They also aim to foster development and **environmental benefits** and **catalyse new sources of finance** for the countries.

We will focus on two of seven evaluation questions, which were most relevant to realist analysis:

- Are investments resulting in the outputs, outcomes, and impacts expected in the ToC?
- To what extent is CP3 leveraging **additional** private and public finance in developing countries?



Activity: Why does CP3 seem well suited to a realist analysis?

Write your evaluation questions & theory of change

RA is appropriate when evaluating...

- ✓ Complex programmes with wider learning potential
 - **YES**
- ✓ Programmes that sit within a system:
 - **YES**
- ✓ Programmes that produce mixed outcomes
 - **POSSIBLY**

RA is not appropriate when...

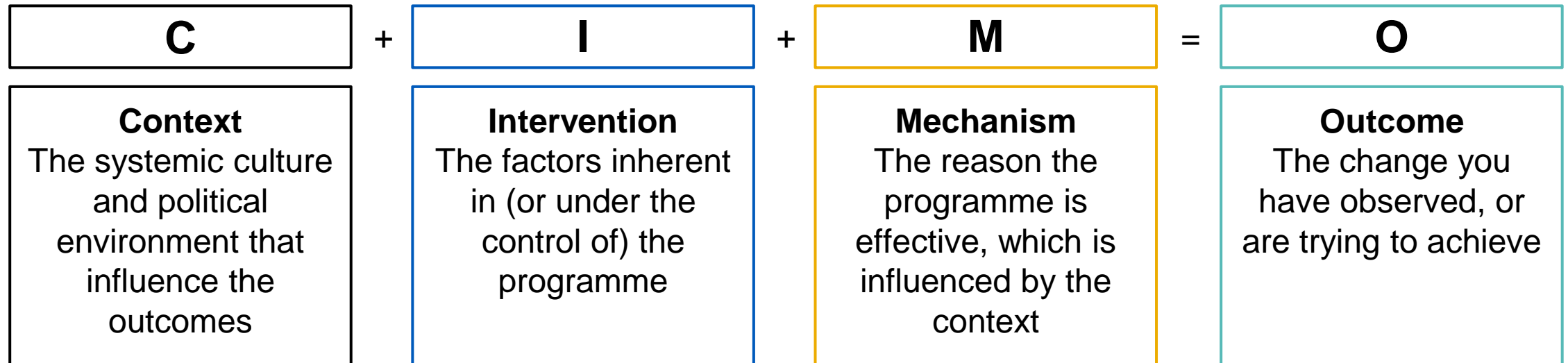
- × The how, why and where programmes work is already understood
 - **NO**
- × You are chiefly interested in the effect of the programme
 - **NO**

*‘...realist synthesis was appropriate as **the context** where individual investments are operating **makes important differences to the outcomes** and that no programme works everywhere, for everyone. The approach used explored why interventions may or may not work, in what contexts and for whom: considering **how contextual factors**, including those at the investment level and at the global level, **have shaped and influenced any outcomes or changes observed.**’*

CP3 Team

Develop CIMO statements from Theory of Change

Realist Analysis hypotheses and results are written as CIMO statements



Gather evidence to investigate, test and refine CIMO statements



To guide synthesis report

- 5 case studies
- Value for money analysis
- Statistical assessments of the investment portfolio
- Assessments of past and current market contexts
- Thematic analysis of programme documentation



To guide medium term evaluation

- Financial and non-financial results reporting
- Interviews
- Review of the literature and data (e.g. on private equity)
- Consultations with key stakeholders (e.g. investors)

Write up your results



Activity: Below are two CIMO statements from our CP3 case study. Which part of each statement refers to each of the four C-I-M-O ingredients?

Anchor Effect – By acting as an anchor investor providing public funds at scale and providing management and governance where there was a lack of other investors willing to act as anchor, public and private investors see UK Govt as a trusted investor and reassess risk and reward and invest in these funds, supporting them to leverage additional finance.

Capacity Building (SCAF) – By providing enterprise development support and seed capital financing to early stage clean energy projects and companies in a context where these projects are financially viable, operate in a supportive enabling environment and understand the SCAF investment standard requirements, SCAF-supported companies are able to develop in a commercially sustainable manner (undertake feasibility studies, demonstrate their commercial viability, ESG compliance, etc.) and, as a result, access conventional finance to reach financial close and flowback the initial SCAF financial support.

Write up your results

Anchor Effect – By acting as an anchor investor providing public funds at scale and providing management and governance (I) where there was a lack of other investors willing to act as anchor (C) public and private investors see UK Govt as a trusted investor and reassess risk and reward (M) and invest in these funds, supporting them to leverage additional finance (O).

Capacity Building (SCAF) – By providing enterprise development support and seed capital financing to early stage clean energy projects and companies (I) in a context where these projects are financially viable, operate in a supportive enabling environment and understand the SCAF investment standard requirements (C) SCAF-supported companies are able to develop in a commercially sustainable manner (undertake feasibility studies, demonstrate their commercial viability, ESG compliance, etc.) (M) and as a result access conventional finance to reach financial close and flowback the initial SCAF financial support (O).

Write up your results



Are investments resulting in the outputs, outcomes, and impacts expected in the CP3 Theory of Change?



Yes. CP3 investments have produced development and environmental impacts that exceeded expectations.



To what extent is CP3 leveraging additional private and public finance in developing countries?



Yes. 71% of investment in funds, projects and companies was from the private sector.

Assess your narrative

The CP3 team...



Held an internal meeting to support sense-making and interpretation of findings following the data collection.



Department for
Business, Energy
& Industrial Strategy



Department
for International
Development

Held workshop with HMG stakeholders across departments (DFID and BEIS).



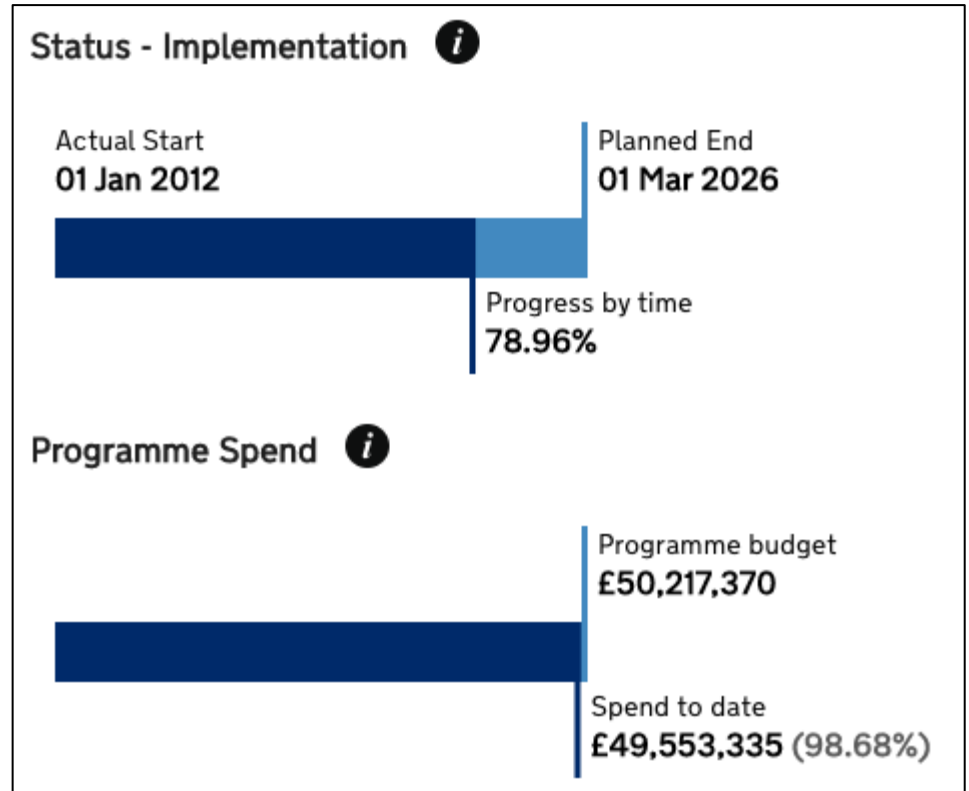
Activity: What kind of other people, processes or information could have been used for a more robust assessment?

Use your research

CP3 continues to be funded.

From the Annual Review, 2022:

This year the methodology used to determine CP3's results has been revised to better reflect the contribution the programme is making at the project-level, in terms of its additionality and the attribution of the overall results achieved.



Contribution Analysis

Fundamentals of Contribution Analysis (CA)



Contribution analysis refers to a process which seeks to identify the ***contribution*** a particular programme made to a change that is observed.



The goal is to provide a plausible explanation that an outside observer would agree with.

When would I use a CA?

CA is appropriate when...

- ✓ You cannot run an RCT or QED
- ✓ There are many different factors that could contribute to change
- ✓ There are many different programmes that can lead to a change

Four main conditions required to infer causality...

1. **Plausibility:** the assumptions behind why the program is expected to work are plausible
2. **Fidelity:** the activities of the programme were implemented
3. **Verified Theory of Change:** the chain of expected results occurred and was verified by evidence
4. **Accounting for other influencing factors:** factors influencing the programme were assessed and did not make a significant contribution or (if they did) the relative contribution was recognised

Contribution Analysis: Coronavirus Community Support Fund (CCSF)

This case study is an evaluation of a DCMS funded intervention, the Coronavirus Community Support Fund (CCSF).

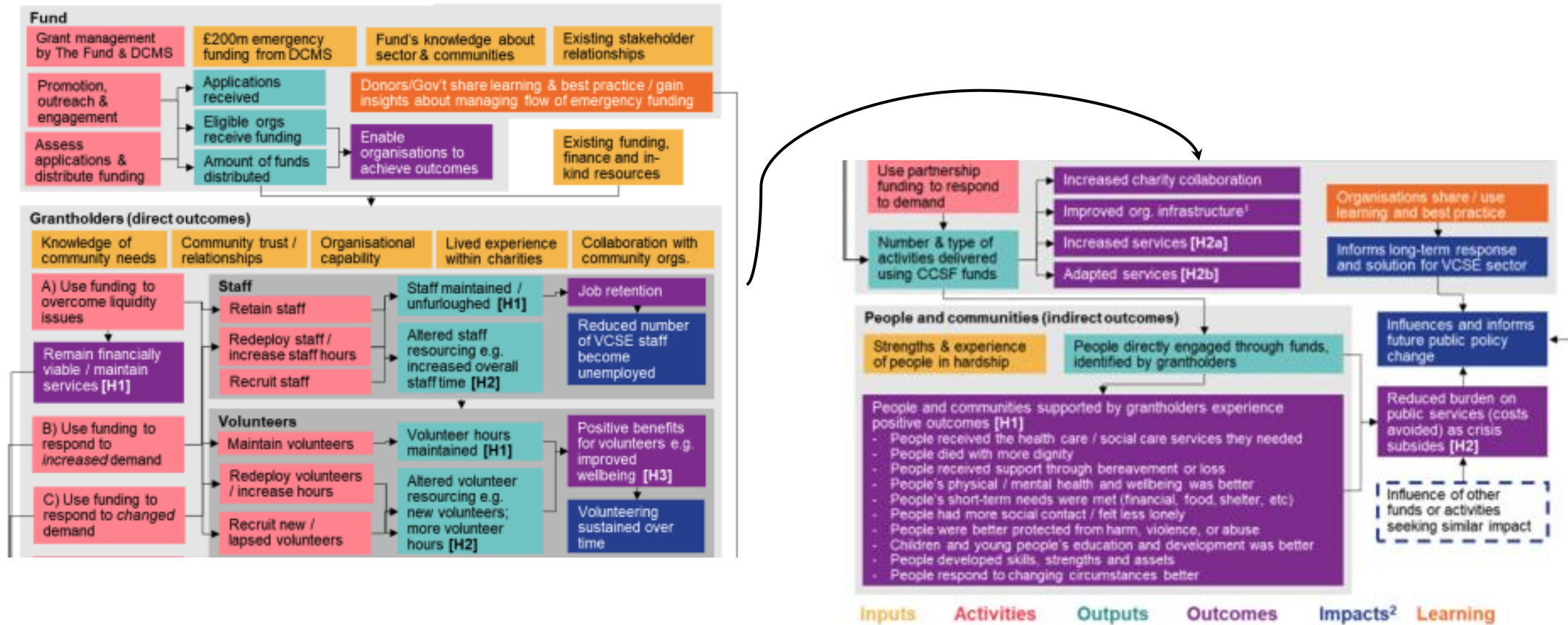
The CCSF was targeted at small and medium sized community organisations delivering activities and support to people affected by the COVID-19 crisis. Between May and October 2020, grants were awarded to **increase support** for vulnerable people affected by the crisis, and to **reduce closures** of essential organisations that provided services for vulnerable people.

The bulk of the impact evaluation was grounded in a **theory-based contribution analysis** in the absence of a robust and data-rich quantitative counterfactual.



Activity: Based on what you have learned about the CCSF goals, what Evaluation questions might you write?

Start with the evaluation questions you want to answer



¹ Organisational infrastructure is specific to the Fund's COVID-19 Emergency Response programme.

² Given the limited evaluation timeframe, only indicative evidence will be available to extrapolate longer term effects using existing evidence, where possible.

Write your evaluation questions & theory of change



The CCSF will increase community support to vulnerable people affected by the COVID-19 crisis, through the work of civil society organisations.



The CCSF will reduce temporary closures of essential charities and social enterprises, ensuring services for vulnerable people impacted by COVID-19 had the financial resources to operate.



Activity: What kind of evidence could the DCMS gather to attribute contribution of the CCSF in this evaluation? Make a list and then we'll discuss.

Gather evidence to investigate your claim

The CCSF team gathered the following evidence:



Two **online surveys** with
grantholders and
volunteers

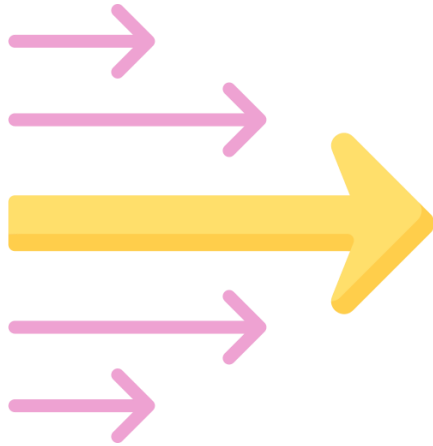


Online interviews with
grantholders







Case studies involving
interviews with grantholder
staff, volunteers, delivery
partners and programme
beneficiaries

Write up your results



A contribution narrative is a written and visual depiction of **how the programme contributed to a change**, including the role of any other programmes or external factors

HYPOTHESIS ASSESSMENT	
<p>Degree of confidence</p> 	<p>CCSF grantees were successful in reaching people and communities disproportionately affected by COVID-19. They were found to have maintained, increased and / or adapted their activities to deliver appropriate support to different groups of beneficiaries and were confident that this had contributed to a range of positive outcomes for people and communities</p> <p>Supporting evidence includes:</p> <ul style="list-style-type: none">• The most common beneficiary groups were people with mental health conditions, people with longstanding illnesses or disabilities, children and young people, and people and families facing financial hardship.• Grantees used a range of approaches to deliver activities and support to beneficiaries, with an estimated 6,400 offering multiple methods of engagement to beneficiaries including phone, video call and face-to-face.• Most grantees reported some of their beneficiaries had increased social contact, better mental health and wellbeing and felt less lonely.• Outcomes relating to reduced loneliness and isolation were referenced by beneficiaries who participated in case studies.
 	
ALTERNATIVE EXPLANATIONS	
	<p>Grantees acknowledged that there were other potential sources of support available to beneficiaries, including other local and statutory services, and that these could also have contributed to positive outcomes. For some short-term outcomes, such as basic needs being met, the contribution of CCSF is clear. For others, the contribution of CCSF is more speculative in the absence of direct measurement and there are likely to have been other contributing factors.</p>

Assess your narrative

The CCSF team...



... used an internal Expert Advisory Board (senior members of contracted organisations) to receive written and verbal feedback on outputs.



... completed Quality Assurance of analysis and reports at the contracted organisations.



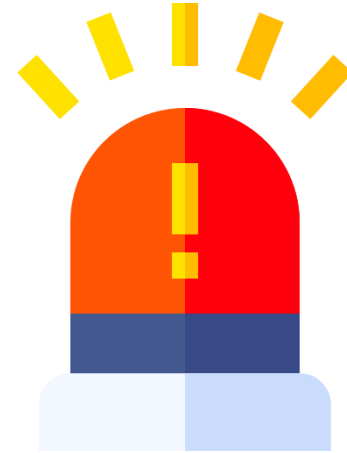
Activity: What kind of other people, processes or information could have been used for a more robust assessment?

- + Robust teams
- Involve stakeholders with no ties to the organisations.

Use your research



The evaluators concluded that they had identified promising evidence of the fund on its two main objectives.



Contributed to the evidence base for deploying and evaluating emergency funding programmes.

Benefits, risks and uses of TBEs

What are the benefits of theory-based designs?



Added rigor



Why or how a programme works



Confidence to apply programme in a new context

Cabinet Office: GREAT Campaign

M/E framework was scrutinised and approved by key bodies.

DLUHC: Supported Housing

Helped support the business case for successor programme.

MoJ: Reforms to HMCTS

Evaluate a large and complex programme, made up of many individual projects.

Scottish Gov: Inward investment

Data rich. Broad coverage of the impact with no counterfactual.

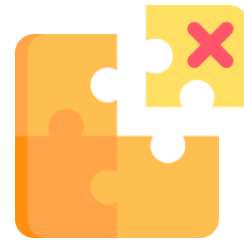
What are the limits and risks of theory-based designs?



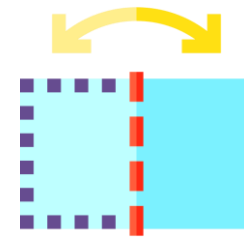
Cost



Time



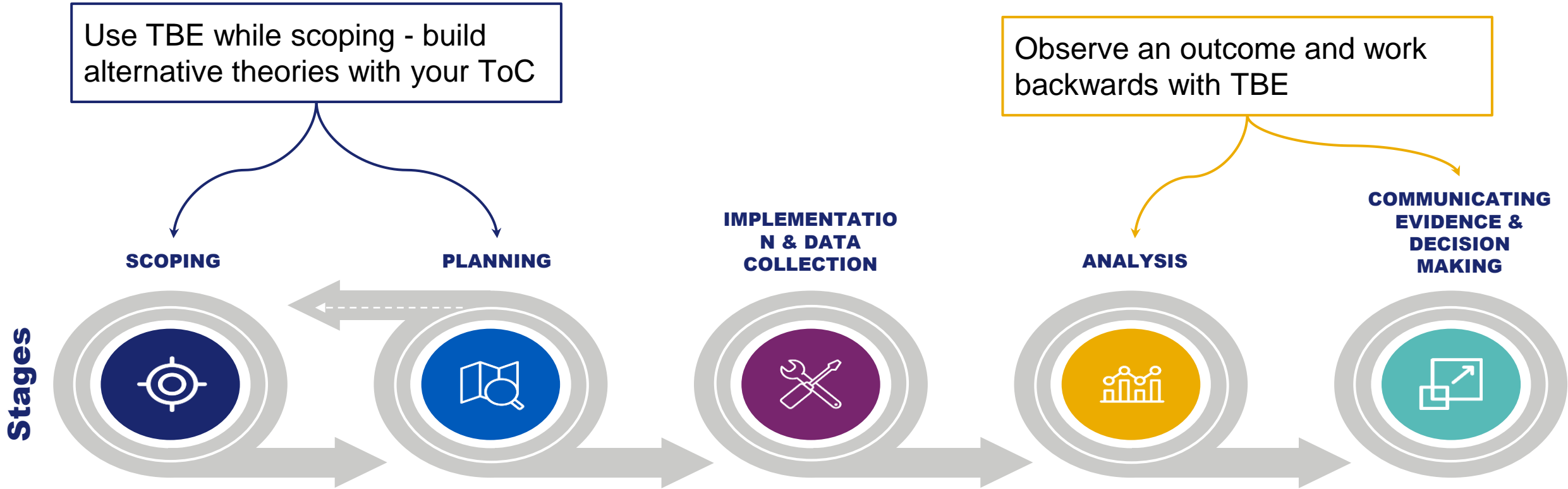
Misapplied



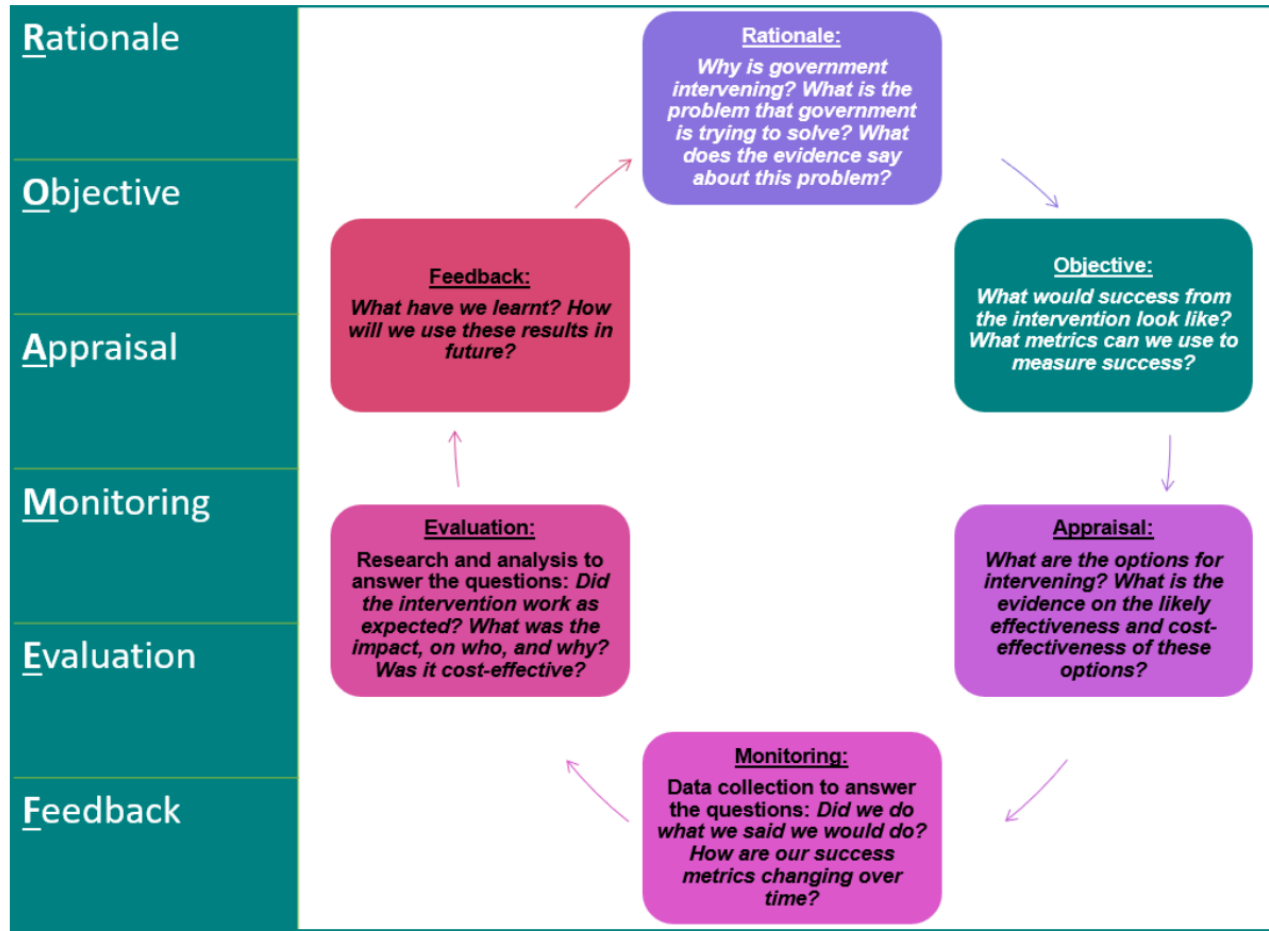
Only useful after a change

Advocacy and application of learning

Advocating for the use of theory-based evaluation



Including TBE across the policy lifecycle



Activity: How does what you have learned today fit into the ROAMEF cycle?

- Think about an upcoming or current evaluation or policy you are involved in. How can you apply your learning from this module to influence that work?
- What barriers exist? How do you push through? What people or resources can support you?
- Write an intention for how you will use this in your work in the next 1-2 months.

Summary

In this module, we have learnt:

- What a TBE is and when to run one (e.g. complex intervention, long timescales, no counterfactual available).
- To identify the types of evaluation questions that TBE designs can and cannot answer.
- How TBE designs build on a ToC and can work alongside other evaluation methods.
- How to run a Realist Analysis and Contribution Analysis.
- To describe other theory-based designs (process tracing and qualitative comparative analysis).
- The key things that need to be considered when preparing and running each TBE design.
- The benefits and risks of different TBE designs.
- Contrast the most appropriate TBE method(s) to use based on the policy context.
- How to critically assess the findings of a TBE and advocate for their use across the policy cycle.

Further resources

Resource
The Magenta Book : Section 3.4. Theory-based impact evaluation methods
The Magenta Book Annex A : Section A1 Theory based methods for impact evaluation
PHE : A brief introduction to realist evaluation
Applying contribution analysis : Lessons from five years of practice
Better Evaluation : Short introductions to many TBE designs
The Green Book
An example of a Realist Analysis