Introduction to Evaluation

Evaluation Task Force Academy 2.0

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

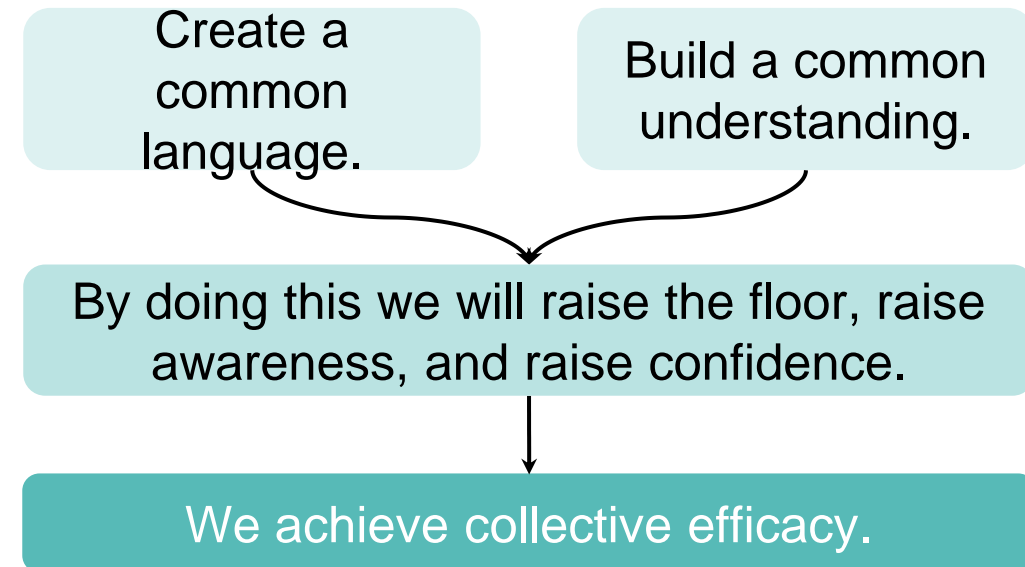**Module 6:** Impact Evaluation - Quasi-Experimental Designs

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.

Create a common language.

Build a common understanding.

By doing this we will raise the floor, raise awareness, and raise confidence.

We achieve collective efficacy.

Module 1: Overview of Content

- [Learning outcomes](#)

- [Why evaluate?](#)

- [Stages of Evaluation](#)

- [Key Evaluation Methods](#)

- [Advocacy and application of learning](#)

Key
Know/Comprehend: lowest level of complexity
Apply/Analyse: medium level of complexity
Evaluate/Create: highest level of complexity

Learning outcomes

I can **explain** what evaluation means and why it is important to HMG

I can **list** and **describe** the key stages of an evaluation

I can **describe** the difference between:

- Process Evaluations,

- Impact Evaluations,

- Value for Money Evaluations

I can **advocate** for including evaluation across the policy cycle

I can **influence** policy design to ensure evaluation methods are used

Why evaluate?

What is an evaluation?

1. **Asking evaluation questions**

2. **Making a plan** to answer them

3. Collecting **data**

4. **Answering** questions

5. **Informing** decisions

We can learn what works..

Do Covid vaccines prevent symptomatic Covid-19 illness?

Did the ONS Covid Infection Survey impact government decision-making?
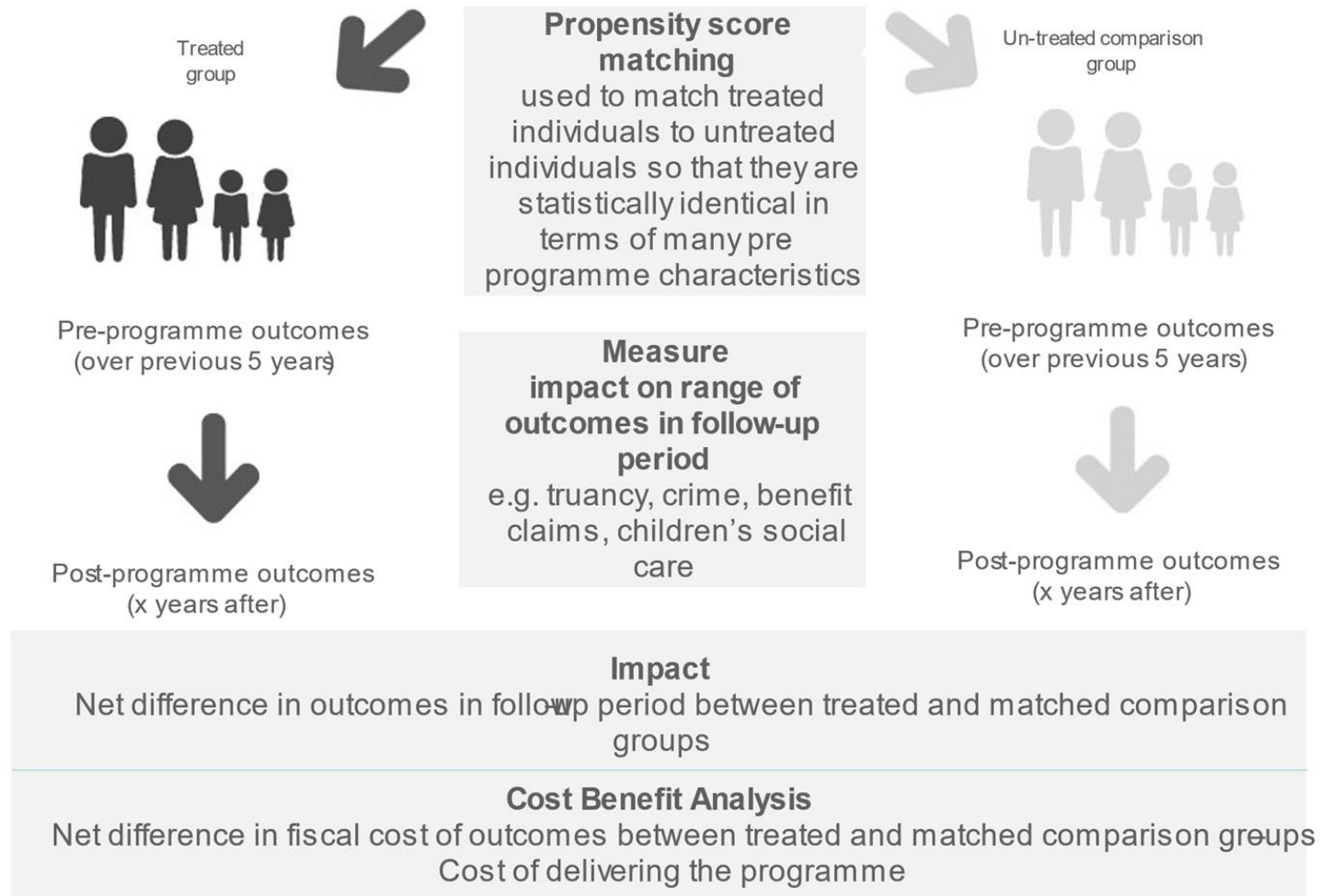
Did the 'Troubled Families' programme improve family outcomes?

Troubled Families Impact Evaluation

Ministry of Housing, Communities & Local Government

How does the Impact Evaluation work?

Treated group

Propensity score matching
used to match treated individuals to untreated individuals so that they are statistically identical in terms of many pre programme characteristics

Un-treated comparison group

Pre-programme outcomes (over previous 5 years)

Pre-programme outcomes (over previous 5 years)

Measure impact on range of outcomes in follow-up period
e.g. truancy, crime, benefit claims, children's social care

Post-programme outcomes (x years after)

Post-programme outcomes (x years after)

Impact
Net difference in outcomes in follow-up period between treated and matched comparison groups

Cost Benefit Analysis
Net difference in fiscal cost of outcomes between treated and matched comparison groups
Cost of delivering the programme

Flexible Working Impact Evaluation

Indeed: Can a timely prompt produce thousands of flexible jobs?

THE BEHAVIOURAL INSIGHTS TEAM | indeed one search. all jobs. | Government Equalities Office

What is the problem?

9 in 10 people want to **work flexibly** but few jobs are advertised as such.

This lack of flexible working options **disproportionately** impacts women who are currently **3x** more likely to work part-time and **2x** more likely to work flexibly.

New **BIT research** has shown that both men and women prefer jobs that offer **clear flexible options,** but few employers remember to include these.

What did we do?

We ran two **randomised control trials (RCTs)** with Indeed, the global job site, to make it easier for employers to **advertise flexible vacancies.**

NEW JOB ROLE
Health Care Assistant
London
Full time
Flexible working options for this role
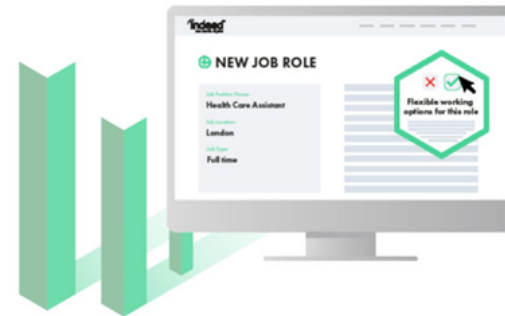
Employers in the treatment group were prompted with a screen that **encouraged** them to think about what **flexible working options they could offer** for the role they were advertising.

More than **150,000 employers** posted over half a million job ads, and almost **20 million individual** job seeker applications passed through this trial making it one of the **largest RCTs ever conducted!**

What happened?

These on-screen prompts led to around a **20% increase of the number flexible vacancies being advertised** on the Indeed site.

Jobs that had flexible working on offer attracted up to **30% more candidates** for the role.

Scaled up an intervention like this could add up to **174,000 additional flexible jobs** to the UK labour market per year.

10

...and learn from what doesn't
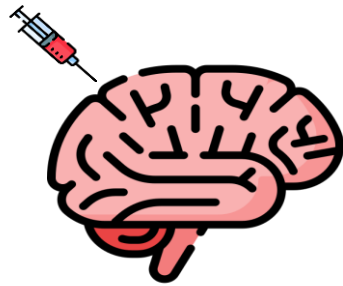
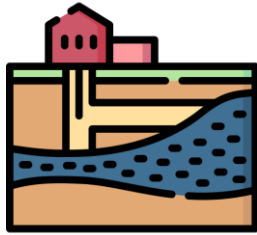**Do steroid injections reduce deaths from head injury?**

Does cognitive-behaviour therapy reduce reoffending among sex offenders?

Do 'playpumps' help people in developing countries to access water?

Evaluations can improve programme delivery

How could a project to map buried assets like cables and pipes avoid scope creep?

When should the National Citizen Service commission its evaluations?

Does expanding a pupil writing intervention that was effective in 26 schools produce the same results when delivered in 167 schools?

Other reasons to evaluate

✓ Protect effective programmes

✓ Hold HMG accountable

✓ Build trust

✓ Minimise bias

✓ Test acceptability of new programmes

✓ Contribute to a broader evidence base

✓ Assess value for money

Common objections to evaluation

We don't need to evaluate - we already know what works

Time and money is better spent on the front-line

We don't have time - the project has already launched!

It looks bad if we find that a programme doesn't work

Activity: What are some common challenges to running evaluations that you have experienced?

Which 'reasons to evaluate' do you think would help persuade particular stakeholders?

Difference between monitoring and evaluation

Monitoring data are collected throughout an intervention to provide answers to a number of policy, research and performance questions. This data typically covers all aspects of an intervention's operation and is generally used to help track progress of an intervention's delivery.

Well designed and implemented **evaluation** provides an understanding of the actual economic, financial, social and environmental impacts of a policy; and/or provides an assessment of how it is/was implemented, why it did or did not deliver as expected, and whether it represents value for money.

Monitoring data <u>alone</u> cannot tell us whether an intervention is working or not.

Stages of evaluation

Stages of evaluation

Stages

SCOPING | **PLANNING** | **IMPLEMENTATION & DATA COLLECTION** | **ANALYSIS** | **COMMUNICATING EVIDENCE & DECISION MAKING**

Magenta book stages

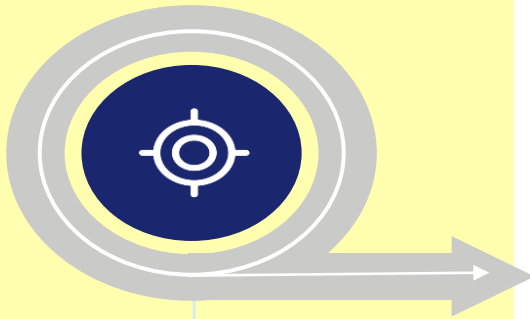| Evaluation Scoping | Conduct the evaluation | Disseminate, use & learning |

Evaluation Design

Choose the appropriate methods

Scoping & Planning stages can occur at the same time and support each other.

Stages of evaluation - Scoping
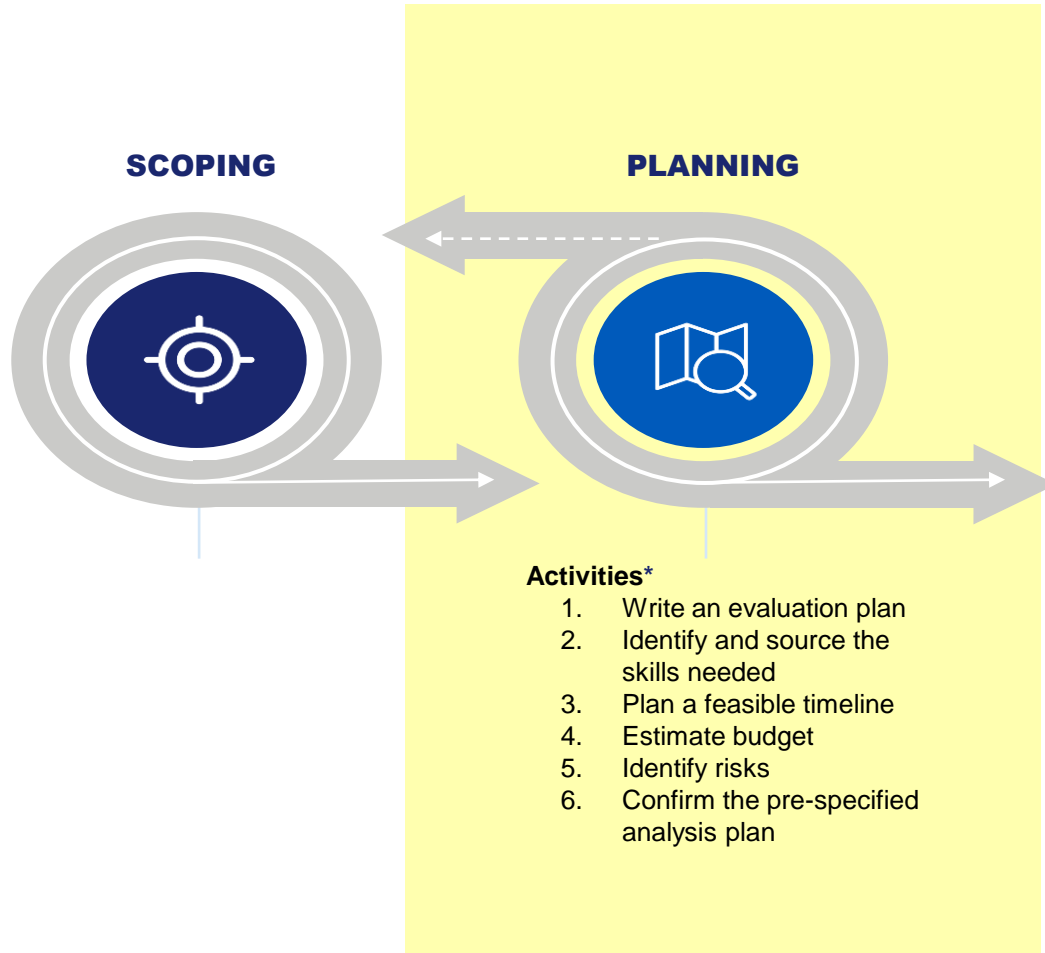
Stages

SCOPING

Activities*
1. Create a Theory of Change
2. Determine the purpose
3. Write evaluation question(s)
4. Define outcome measures
5. Identify primary intended users
6. Choose an evaluation method

Scoping & Planning stages can occur at the same time and support each other.

Stages of evaluation - Planning

SCOPING

PLANNING

Activities*
1. Write an evaluation plan
2. Identify and source the skills needed
3. Plan a feasible timeline
4. Estimate budget
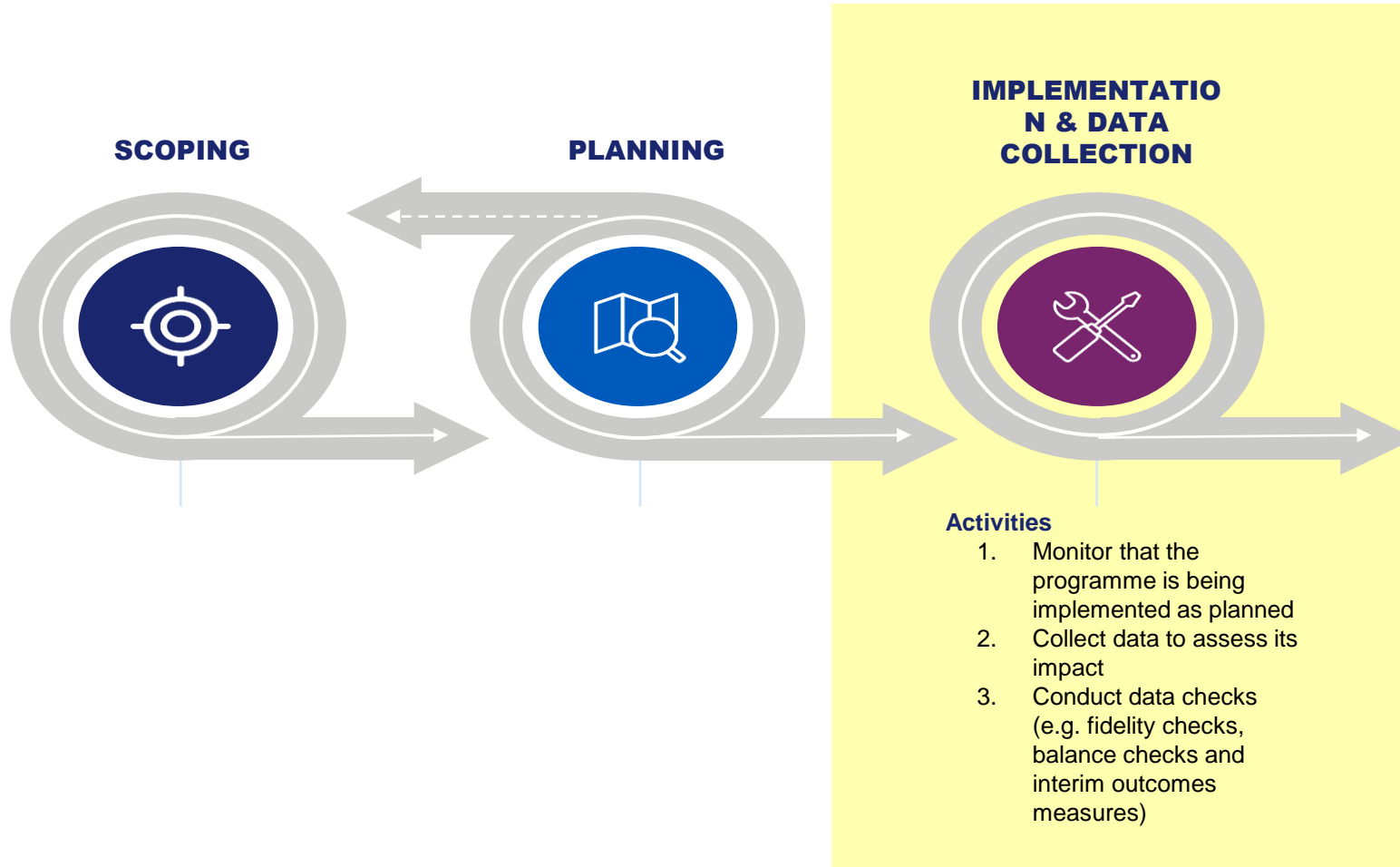5. Identify risks
6. Confirm the pre-specified analysis plan

Scoping & Planning stages can occur at the same time and support each other.

Stages of evaluation - Implementation & Data Collection

Stages

SCOPING

PLANNING

IMPLEMENTATION & DATA COLLECTION

Activities
1. Monitor that the programme is being implemented as planned
2. Collect data to assess its impact
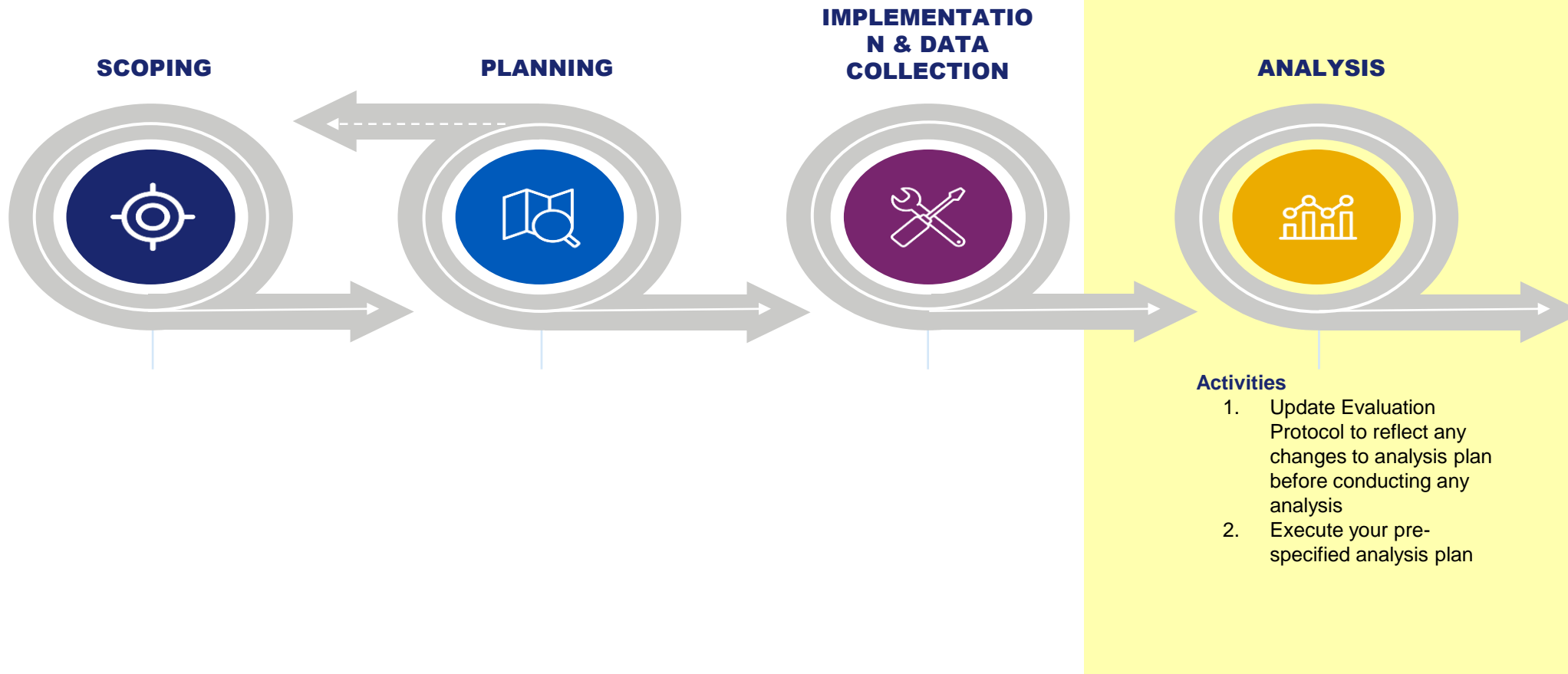3. Conduct data checks (e.g. fidelity checks, balance checks and interim outcomes measures)

Scoping & Planning stages can occur at the same time and support each other.

Stages of evaluation - Analysis

Stages

SCOPING

PLANNING

IMPLEMENTATION & DATA COLLECTION

ANALYSIS

Activities

1. Update Evaluation Protocol to reflect any changes to analysis plan before conducting any analysis
2. Execute your pre-specified analysis plan

Scoping & Planning stages can occur at the same time and support each other.

Stages of evaluation - Communicating Evidence & Decision Making

Stages

SCOPING

PLANNING

IMPLEMENTATION & DATA COLLECTION

ANALYSIS

COMMUNICATING EVIDENCE & DECISION MAKING

Activities
1. Produce technical outputs
2. Produce outputs for a non-technical audience
3. Engage key stakeholders

Scoping & Planning stages can occur at the same time and support each other.
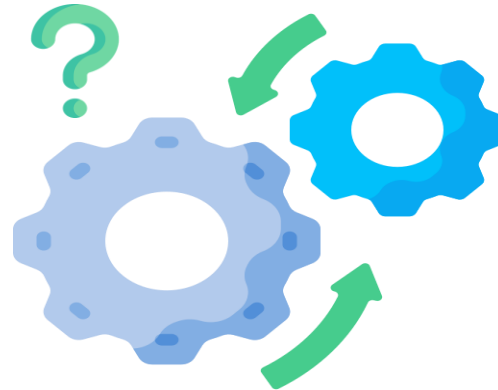
Key evaluation methods

Choosing an evaluation method

Before we can choose a method that works for our project, we first have to determine which category of evaluation our evaluation questions fit into.

Impact questions   OR   **Process questions**   OR   **Value-for-money questions**

Impact questions

- Focused on **outcomes**
- Tell you **the effect of something**
- Make a **comparison**

Example questions
- What effect is my programme or policy having?
- Are the people that the programme or policy is supposed to serve better off because of it?
- Do more people sign up for my programme if I change the recruitment materials?

Pros
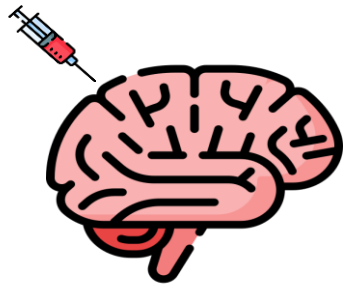- ✓ Tell us whether a programme is effective or not
- ✓ More generalisable results

Cons
- × (Quantitative impact evaluation) can't explain *why* a programme does or doesn't work

These are all impact questions!

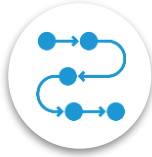Do steroid injections reduce deaths from head injury?

Does cognitive-behaviour therapy reduce reoffending among sex offenders?

Do 'playpumps' help people in developing countries to access water?

Process questions

- *What's* happening and *how*
- **Understanding programme outputs***

Example questions
- How is my programme currently operating? Does this match what was intended?
- Who is the programme or policy serving? Is this the intended audience?
- How do participants interact with the programme or policy? Is there any place they fall off?
- How much staff time is going into running the programme?

Pros
- ✓ Understand the range and diversity of experiences
- ✓ Tell us how prevalent or common an experience or barrier is
- ✓ Tell us why a programme worked (or didn't)
- ✓ More flexible and can adapt to suit changing circumstances

Cons
- ✕ Can't tell us whether something worked

- Monitoring is sometimes referred to as process evaluation, because it focuses on the implementation process of a project. However, it only tracks progress towards programme goals and is not an evaluation method as it does not measure the extent to which these changes that are monitored are contributing to the outputs of a program.

Choosing an evaluation method: impact vs. process

Activity: Which type of evaluation is best suited to answer the following questions:

1 Does exporting affect the profits and productivity of rug manufacturers in Egypt?

2 Does a text message increase the uptake of Covid-19 vaccinations?

3 How do people experiencing unemployment engage with employment support from DWP?

Value-for-Money questions

- Focused on **cost effectiveness**
- Do the benefits of a programme justify the costs?

Example questions
- To what extent did the chosen delivery provide cost efficiencies compared to other options?
- Are we buying inputs of the appropriate quality at the right price?
- How well are the outputs produced by an intervention having the intended effect?

Pros
- ✓ Offer a comprehensive evaluation framework
- ✓ Ensure that funds are used where they can have the greatest impact
- ✓ Increase accountability and transparency in how public funds are used

Cons
- ✗ Risk of oversimplifying complex programmes by reducing them to quantitative measures
- ✗ Negative impact on innovation as novel approaches come with higher uncertainties
- ✗ Risk of lacking stakeholder buy-in

Evaluation and Benefits Realisation

Benefits realisation

Evaluation

Benefits…
- identification
- quantification and analysis
- planning and tracking
- realisation and optimisation

Benefits & Evaluation…
- defining success
- identifying metrics
- tracking and monitoring
- quantification and analysis
- feedback to delivery

What works, what doesn't and why?
- what difference it made (*impact* evaluation)
- was it worth it (*value for money* evaluation)
- how it was delivered (*process* evaluation).

Many evaluations use multiple methods

Case study: Covid Job Retention Scheme

Process
What was the experience of those who used the Coronavirus Job Retention Scheme?

Impact
To what extent did the Coronavirus Job Retention Scheme protect jobs?

Benefits Management
Did the scheme provide value for money?

16 March 2020
PM announced people should work from home *"where they possibly can".*

20 March 2020
Chancellor announced the CJRS (claims from 1 March)

23 March 2020
PM announced lockdown measures ('Stay Home, Protect the NHS, Save Lives')

20 April 2020
The CJRS system went live

12 May 2020
Chancellor announced the CJRS Flexible Furlough

1 July 2020
The CJRS Flexible Furlough began

12 October 2020
PM introduced local COVID Alert Levels

31 October 2020
PM announced lockdown and the CJRS Extension

1 November 2020
The CJRS Extension began

5 November 2020
Chancellor announced further plans to extend the CJRS Extension

Second lockdown began (in England)*

Advocacy and application of learning

How can we advocate for the use of evaluation?

Activity: Imagine that you work for the department that is about to launch Photo ID requirements for an election in May 2023:

- How might you convince a policy team member of the need to evaluate the programme?
- Which of the reasons included in this session do you think would be most persuasive?
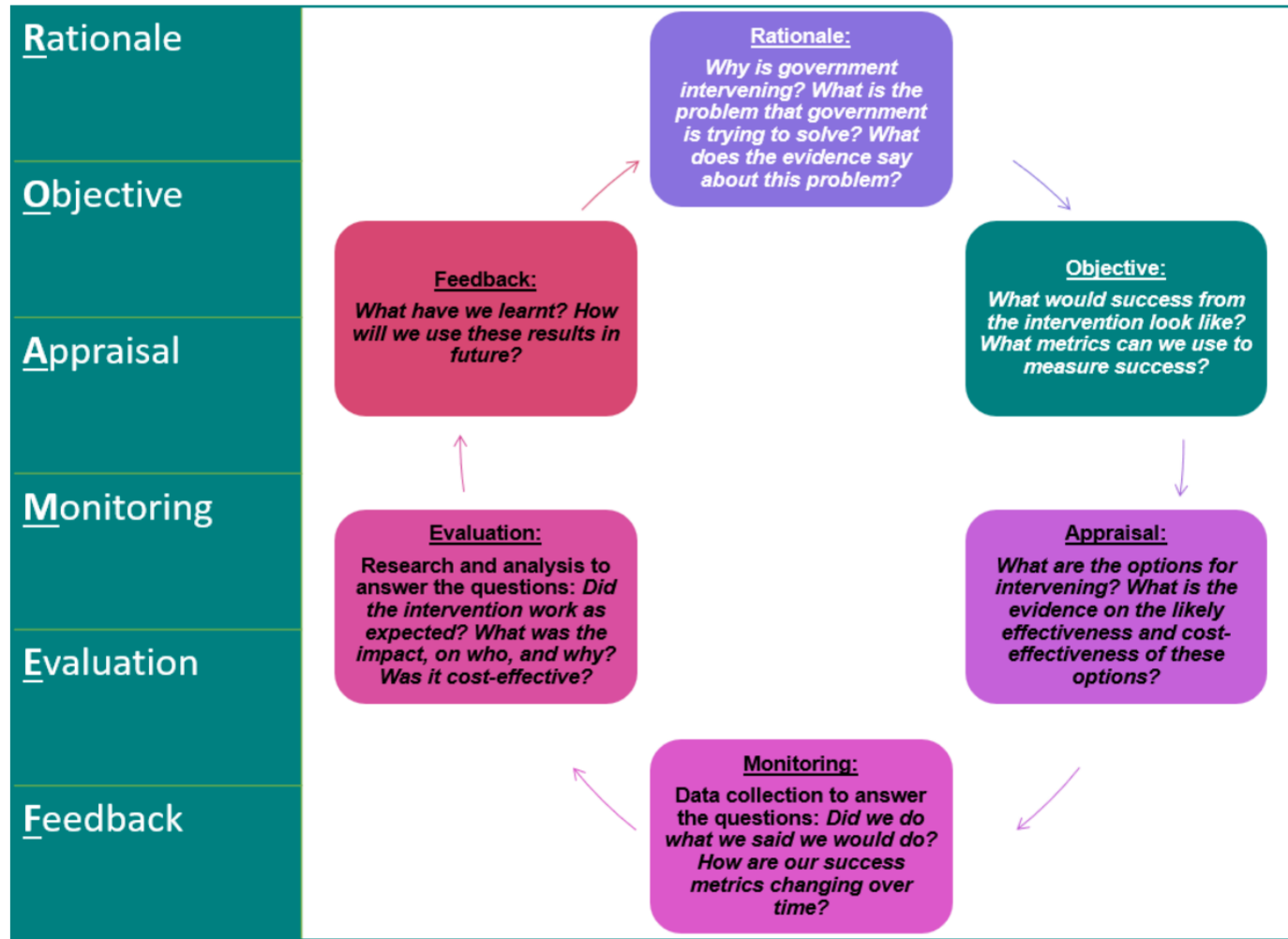
Including evaluation across the policy lifecycle

Rationale
Objective
Appraisal
Monitoring
Evaluation
Feedback

Rationale:
Why is government intervening? What is the problem that government is trying to solve? What does the evidence say about this problem?

Objective:
What would success from the intervention look like? What metrics can we use to measure success?

Appraisal:
What are the options for intervening? What is the evidence on the likely effectiveness and cost-effectiveness of these options?

Monitoring:
Data collection to answer the questions: Did we do what we said we would do? How are our success metrics changing over time?

Evaluation:
Research and analysis to answer the questions: Did the intervention work as expected? What was the impact, on who, and why? Was it cost-effective?

Feedback:
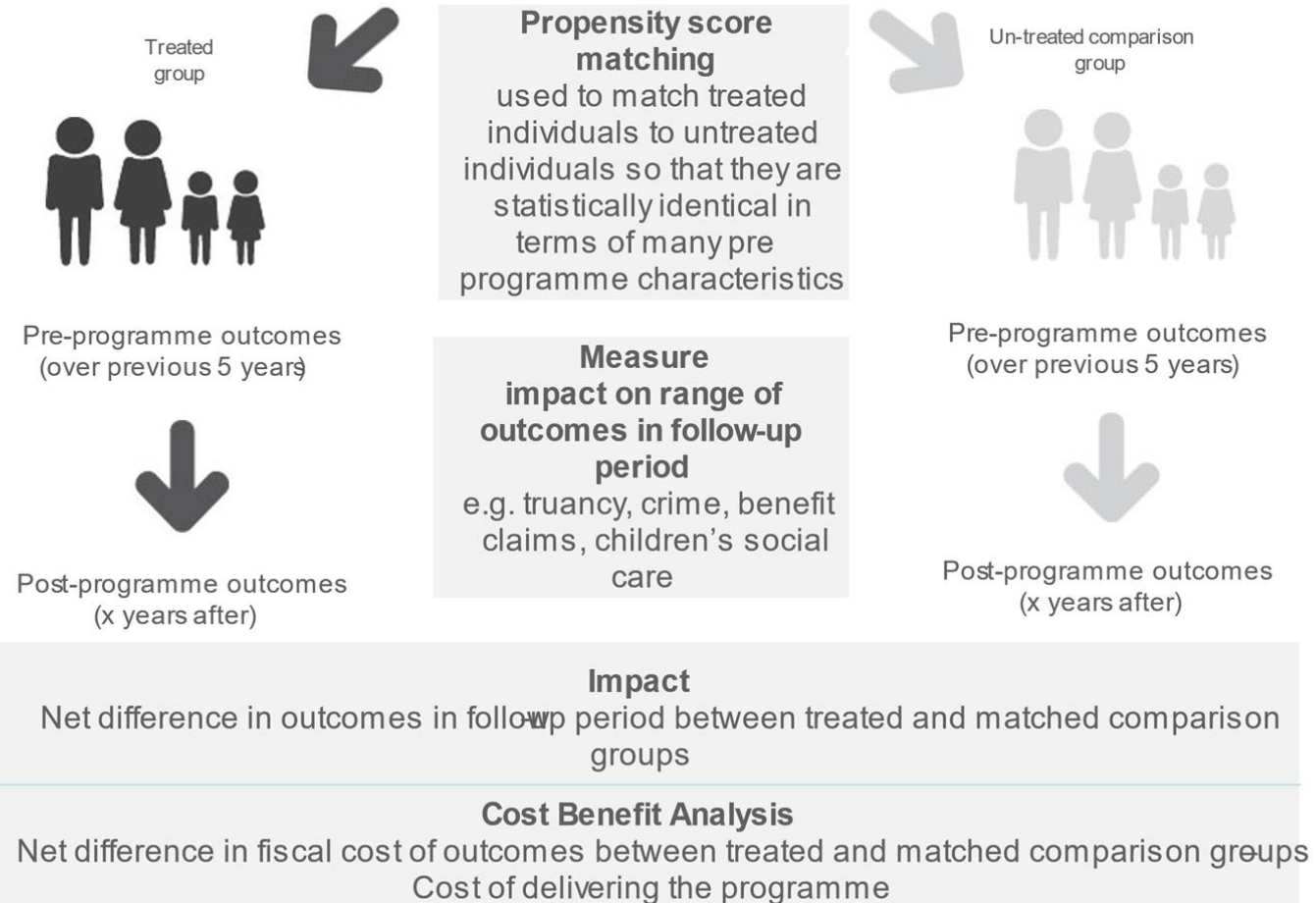What have we learnt? How will we use these results in future?

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

Troubled Families Impact Evaluation

Ministry of Housing, Communities & Local Government

How does the Impact Evaluation work?

Treated group

Propensity score matching
used to match treated individuals to untreated individuals so that they are statistically identical in terms of many pre programme characteristics

Un-treated comparison group

Pre-programme outcomes (over previous 5 years)

Measure impact on range of outcomes in follow-up period
e.g. truancy, crime, benefit claims, children's social care

Pre-programme outcomes (over previous 5 years)

Post-programme outcomes (x years after)

Post-programme outcomes (x years after)

Impact
Net difference in outcomes in follow-up period between treated and matched comparison groups

Cost Benefit Analysis
Net difference in fiscal cost of outcomes between treated and matched comparison groups – Cost of delivering the programme

Putting your learning into practice

Activity:

- 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 evaluation means and why it is important to HMG.
- How monitoring and evaluation differ, common objections to evaluation and how to tackle these.
- The key stages of evaluation: i) scoping; ii) planning; iii) implementation & data collection; iv) analysis; v) communicating evidence and decision making.
- The main types of evaluation and key differences between:
  - **Process Evaluations,**
  - **Impact Evaluations,** and
  - **Value for Money Evaluations**.
- The difference between Evaluation and Benefits Realisation.
- How evaluation should be included across the policy cycle and how to influence policy design to ensure evaluation methods are used.

Further resources

Resource
Evaluation strategies for each government department
Evaluation and Trial Advice Panel
The Magenta Book: Central Government guidance on evaluation
Guidance on benefits management
ETF: Resources for evaluating policy in government
BIT: TESTS
The Green Book