Evaluating OD: The Diamond Standard

Dr Dione Hills
Principal Researcher Consultant
Tavistock Institute of Human Relations
Quality in evaluation: the gold standard

‘Strong’ and ‘Weak’ standards of proof
(HM Treasury and Defra: ‘Quality in Policy Evaluation’)

Strong

• Random allocation/ experimental design
• Quasi experimental designs
• Intervention group vs unmatched comparison group
• Predicted vs actual
• No comparison group

Weak
Quality in evaluation: a diamond standard

Valuable

Multi faceted

Durable

Shiny and attractive
Ten ways evaluation adds value to OD*

1. Talking about evaluation helps to clarify desired outcomes and informs choice and design of interventions
2. Evaluation during an OD intervention helps keep it on track
3. Evaluation can be a valuable OD intervention in its own right
4. Evaluation enables learning
5. The process of evaluation enhances relationships
6. Evaluation helps develop OD as a discipline
7. Evaluation can demonstrate investment in OD is worth while
8. Evaluation feedback can be used to recognise and celebrate change efforts
9. Evaluation can help practitioners with their own professional development
10. Evaluation can help external consultants win business

*Best practice in OD evaluation: Liz Finney and Carol Jefkins
Roffey Park
Putting ‘value’ into evaluation

A valuable evaluation is:
• Useful – answering specific questions asked by specific people (stakeholders)
• Engaging
• Proportionate
• Well planned

Key questions to ask when planning an evaluation
• **Who** will be using the results and what is it that they will want to know? (stakeholder engagement)
• **What** will they be using the results for (key decisions)?
• **When** do they need the results?
• **What** kind of data is most useful?
• **How** to present the results to ensure they are useful?

1. Defining the policy objectives and intended outcomes
2. Considering implications of policy design for evaluation feasibility
3. Defining the audience for the evaluation
4. Identifying the evaluation objectives and research questions
5. Selecting the evaluation approach
6. Identifying the data requirements
7. Identifying the necessary resources and governance arrangements
8. Conducting the evaluation
9. Using and disseminating the findings
Inserting evaluation into the programme cycle

Action research cycle

Evaluation = feedback + reflection + data

ROAMEF policy cycle (Green book)
Multifaceted: When to evaluate? Before, during or after?

Before (Appraisal, Ex-anti)

During (Process, learning, developmental)

After (ex anti, impact, outcome)
Multifaceted: Why evaluate?
Different purposes

- Accountability
- Learning
- Obtaining funding
- Knowledge evidence
Multifaceted: How to evaluate?
Different types of evaluation

- Process/formative
- Economic
- Outcome or impact
### Different types of evaluation

<table>
<thead>
<tr>
<th>Type of evaluation</th>
<th>When used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process or formative evaluation</td>
<td>For <strong>learning</strong> and improving, delivery is central aim of evaluation.</td>
</tr>
<tr>
<td>Outcome or impact evaluation</td>
<td>When <strong>accountability</strong> and producing <strong>evidence</strong> of ‘what works’ is important</td>
</tr>
<tr>
<td>Economic evaluation</td>
<td>When <strong>accountability</strong> and making sure that resources are being effectively targeted is important</td>
</tr>
</tbody>
</table>
Multifaceted: How to evaluate?
Different evaluation designs
Experimental designs

population

RCT = Random allocation

Intervention group or area

Counterfactual (control or comparison group or area)

Measurement time 1

Measurement time 2

Measurement time 1

Measurement time 2

Intervention

*Taken from: Quality in policy impact evaluation (HM Treasury, DECC and DEFRA)
Simple outcome and experimental methods work well when:

- Time scale is short to medium term
- Causal pathways are short and straightforward
- The intervention won’t change during implementation
- Context stable and won’t influence outcomes

**Simple outcome approaches** work if the level of change is expected to be large and easy to measure

**Experimental methods** useful when change will be relatively small and require careful measurement
But what if:

- Intervention is developmental or exploratory
- The context is dynamic and changeable
- Time scale is medium or long-term
- Causal pathways are complex or indirect
- Implementation is likely to vary over time
- Impacts uncertain

Use:

- Theory based evaluation methods
- Developmental evaluation: Quinn Patton
  
  ‘Development evaluation is particularly suited to innovation, radical program re-design, replication, complex issues, crises

  In these situations, DE can help by: framing concepts, test quick iterations, tracking developments, surfacing issues.’

  http://betterevaluation.org/plan/approach/developmental_evaluation
Theory based evaluation designs

**Theory of change evaluation**: asks *why* as well as *whether* something worked?...by
- Explores underlying assumptions or theories about why it is working

**Realistic evaluation**: Asks ‘what works for whom, and where’?....by
- Exploring the ‘mechanisms’ by which an intervention brings about change, in a particular context or setting

**Contribution analysis**: asks what *contribution* the intervention made to change? ....by
- Testing alternative explanations for the change taking place

- **All involve mapping the intervention logic**
Theory based evaluations generally require mapping of intervention logic

Step 1: What is the problem?
- Issues being addressed

Step 3: Describe stages between
- Inputs: resources and activities
  - Outputs: activities, participants infrastructure
  - Outcomes: Changes in attitudes, knowledge or behaviour
  - Impacts: Achievement of overall aims

Step 2: What will be the result?
- What are the assumptions about how one step leads to the next one?
Logic mapping helps in visualising intervention as part of a wider system.
Evaluation of a coaching for health staff development activity

**Theory of Change map**

**Inputs**
- Training content
- Delivery style
- Follow up support
- Organisation

**Outputs**
- Delegate engagement
- Satisfaction
- Understanding
- Skill development

**Outcomes**
- Incorporation of coaching skills
- Practitioner well-being
- Better engagement with patients

**Impacts**
- Patient view of self-management
- Patient concordance
- Achievement of health goals

**Kirkpatrick Evaluation Model**

**Level 1 Feedback**
- Follow up survey of delegates

**Level 2 Learning**
- Delegates' evaluation

**Level 3 Behaviour change**
- Follow up conference calls
- Delegates' reflections

**Level 4 Service change**
- Feedback collected from patients

**Possible Data sources**
- Information from trainers, delegates and programme team, observation, training materials,
- Applications/ delegate data
- Feedback forms, interviews, observation,
Shiny and attractive:
Using data that communicates to your stakeholders
Using a range of different data sources

• **Existing data sources:** administrative and monitoring data, large population sets and existing surveys

• **Qualitative methods:** interviews, focus groups, observation, document analysis.

• **Quantitative methods:** numerical data used in statistical analysis. Used extensively in experimental research design and surveys.

• **Creative and indirect methods** including increasing use of social media data
Logic mapping helps identify evaluation questions and data sources

Issues: being addressed

Inputs: resources and activities

Outputs: activities, participants infrastructure

Outcomes: Changes in attitudes, knowledge or behaviour

Impacts: Achievement of overall aims

Background and baseline data may come from policy documents, prior evaluations or appraisal work.

Data on delivery may come from monitoring data or new data collection methods.

Long term outcomes often require new data collection methods, in addition to monitoring or population level data.

© TIHR
Creative and indirect methods

- **Photos, videos**: to illustrate points – may be collected by programme participants

- **Creative ways of gathering participant feedback** (lego pieces, post it notes, reflective discussion)

- **Social media**: can be used to promote discussion, track trends and communicate information

- **Combinations** of all of above (triangulation of data)
Shiny and attractive: communicating results effectively

Think about the formats that suits your audience
- Workshop or verbal presentation (at which results are discussed in detail)
- Summary report or information sheet
- Cartoons, video, webcasts (visual presentations)
- Written report (may or may not be published)
- Journal article (makes results easily accessible for systematic reviews)

Useful points to consider:
- Length
- Language (technical, plain English)
- One report or several (for different audiences?)
- Future accessibility – how people will find it?
Durable: evaluation embedded as a regular activity

- Building in regular cycles of change and reflection
- Select techniques and tools which have been previously effective and are change responsive;
- Streamline evaluation and focus on key priorities and future development directions
- Incorporate into strategic policy and operational activity of service
- Get agreement and commitment of stakeholders
Any questions?
Thank you!
Go sparkle!