Evaluating in complexity: Nearly 20 years on what have we learned?

Judy Oakden: Director, Pragmatica Limited – a member of the Kinnect Group
What we will cover today

01 What grappling with?

02 Two approaches which aid evaluation in complexity

03 Challenges for our practice going forward
You will come away with...

- Two systems approaches:
  - Generic rubric
  - Pattern spotting tool

- Knowing how and why these tools are useful

- A brief update: Eoyang & Berkas 1999
What are we grappling with?

Which translates to this...
Generic rubrics aid evaluation in complex and uncertain environments
Simple rules

- Alignment
- Cohesion
- Separation
Evaluation equivalent is generic rubric

- Flexible
- Adapt to change
- Relevant to context

To address merit, worth and significance in uncertainty
Three parts to a generic rubric

1. Evaluation criteria or aspects of performance
2. Generic performance rating
3. Levels of importance
Case study
1. Developing evaluative criteria
Evaluating in uncertainty

- New Act implemented
- To understand implementation progress
- Stakeholders’ perception of progress
- Many aspects
5.1 Waste levy is imposed/funds collected from landfill operators

5.2 Levy funds allocated to Councils to promote or achieve waste minimisation

5.3 Levy funds are also allocated to contestable projects

5.4 Product Stewardship schemes are accredited

5.5 Councils develop waste minimisation and management plans

5.6 Other – Ministry administration systems, guidance & enforcement, WAB, Levy Review, data collection

5.1 – 5.6 Policy interventions

Waste Minimisation is achieved

Protects environment

Range of benefits ensue

Effective waste minimisation systems and behaviours

Less waste
Recycle more
Manage harmful waste
Responsible producers
Effective systems for waste
Backed by enforcement systems

Case study

Policy interventions

OUTCOMES (END) & IMPACTS

Waste Minimisation is achieved

Protects environment

Range of benefits ensue

Effective waste minimisation systems and behaviours

Less waste
Recycle more
Manage harmful waste
Responsible producers
Effective systems for waste
Backed by enforcement systems

Policy interventions

Case study
In a nutshell

- New business systems
- Processes for funding
- Working in different ways
- Different expectations (operational & planning roles different from in the past)
- New reporting and evaluation roles
5.1 Waste levy is imposed/ funds collected from landfill operators

5.2 Levy funds are allocated to Councils to promote or achieve waste minimisation

5.3 Levy funds are also allocated to contestable projects

5.4 Product Stewardship schemes are accredited

5.5 Councils develop waste minimisation and management plans

5.6 Other - Ministry administration systems, guidance & enforcement, WAB, Levy Review, data collection

Focus - **systemic**, rather than within policy interventions

**Policy interventions**

Waste Minimisation is achieved

- Protects environment
- Range of benefits ensue

Effective waste minimisation systems and behaviours

- Range of benefits ensue
Case study

Basic ‘rules’ underpinning implementation

- Share information to build awareness and compliance
- Efficient administrative processes in operation
- Effective relationship building to support collaboration in the sector
- Use of good practice to build capability and capacity (including infrastructure) across the sector
Could be expressed as ‘simple rules’

**Evaluative criteria**
- Information to build awareness and compliance
- Administrative efficiency of the Act (Ministry’s performance)
- Effective relationships – collaboration in the sector
- Good practice – building capability and capacity (including infrastructure) across the sector

**Simple rules**
- Share information that builds awareness and compliance
- Administer efficiently
- Create and sustain collaborative relationships
- Build capability and capacity to minimise waste

2. Generic performance rating
### Generic performance rating

<table>
<thead>
<tr>
<th>Rating level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent (Always)</td>
<td>• Clear example of exemplary performance or very good practice in this domain: no weaknesses.</td>
</tr>
<tr>
<td>Very good (Almost always)</td>
<td>• Very good to excellent performance on virtually all aspects; while strong overall not exemplary; no weaknesses of any real consequence.</td>
</tr>
<tr>
<td>Good (Mostly, with some exceptions)</td>
<td>• Reasonably good performance overall; might have a few slight weaknesses, but nothing serious.</td>
</tr>
<tr>
<td>Emerging: (Sometimes, with quite a few exceptions)</td>
<td>• Fair performance, some serious, but non-fatal weaknesses on a few aspects.</td>
</tr>
<tr>
<td>Not yet emerging: (Barely or not at all)</td>
<td>• No clear evidence has yet emerged that the aspect of performance has taken effect.</td>
</tr>
<tr>
<td>Poor: Never (Or occasionally with clear weakness evident)</td>
<td>• Clear evidence of unsatisfactory functioning; serious weaknesses across the board or on crucial aspects.</td>
</tr>
</tbody>
</table>

## Case study

### Data collected mapped to evaluative criteria

<table>
<thead>
<tr>
<th></th>
<th>Overall rating</th>
<th>Online survey</th>
<th>Stakeholder focus groups</th>
<th>Individual in-depth interviews</th>
<th>Administrative data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrative efficiency</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Effective relationships</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Good practice</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Information, awareness and compliance</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>
## Data collected mapped to evaluative criteria
(Simplified examples not actual data)

<table>
<thead>
<tr>
<th></th>
<th>Overall rating</th>
<th>Online survey</th>
<th>Stakeholder focus groups</th>
<th>Individual in-depth interviews</th>
<th>Administrative data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrative efficiency</td>
<td>G</td>
<td></td>
<td>G</td>
<td></td>
<td>G</td>
</tr>
<tr>
<td>Effective relationships</td>
<td>G, Em</td>
<td>NY</td>
<td>NY</td>
<td>Em</td>
<td></td>
</tr>
<tr>
<td>Good practice</td>
<td>NY</td>
<td>NY</td>
<td>Em</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information, awareness and compliance</td>
<td>Em, G</td>
<td></td>
<td>Em</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. Levels of importance
When undertaking data synthesis...

- What’s more important?
- How to capture this?
- Deal breakers?

All aspects of performance are not created equal – if we treat them as equal this leads to judgments that don’t give us the best result.
What’s important may change over time

Introduction
- Implement Act
- People know about it and start changing behaviour
- Administrative processes set up

Growth
- Focus on developing relationships
- Collaboration starts to occur
- Administrative processes honed

Maturity
- Learn what best practice looks like and start to document this

Decline
- Aspects of process no longer fit for function – amendments adjustments made

Case study
## Case study

### Example

<table>
<thead>
<tr>
<th>Evaluation of stakeholder perceptions of implementation of Act</th>
<th>Level of importance during the evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>At start</td>
</tr>
<tr>
<td>Administrative efficiency</td>
<td>Moderate</td>
</tr>
<tr>
<td>Effective relationships</td>
<td>Moderate</td>
</tr>
<tr>
<td>Good practice</td>
<td>Low</td>
</tr>
<tr>
<td>Information, awareness and compliance</td>
<td>High</td>
</tr>
</tbody>
</table>
# Case study

How we made the judgements: pattern spotters

<table>
<thead>
<tr>
<th>Generalisation:</th>
<th>In general I noticed ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exception</td>
<td>In general...except...</td>
</tr>
<tr>
<td>Contradiction</td>
<td>On the one hand...but on the other hand...</td>
</tr>
<tr>
<td>Surprise</td>
<td>I was surprised by...</td>
</tr>
<tr>
<td>Puzzle</td>
<td>I wonder...</td>
</tr>
</tbody>
</table>

Adaptation of Pattern Spotters from HSD Wiki [http://wiki.hsdinstitute.org/pattern_spotters](http://wiki.hsdinstitute.org/pattern_spotters)

Case study

Example of making overall judgements  (Simplified examples not actual data)

<table>
<thead>
<tr>
<th>Administrative efficiency</th>
<th>Overall rating</th>
<th>Online survey</th>
<th>Stakeholder focus groups</th>
<th>Individual in-depth interviews</th>
<th>Administrative data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrative efficiency</td>
<td>G</td>
<td>G</td>
<td></td>
<td></td>
<td>G</td>
</tr>
<tr>
<td>Effective relationships</td>
<td>Em</td>
<td>G</td>
<td>Em</td>
<td>NY</td>
<td></td>
</tr>
<tr>
<td>Good practice</td>
<td>NY</td>
<td>NY*</td>
<td>Em</td>
<td>Em</td>
<td></td>
</tr>
<tr>
<td>Information, awareness and compliance</td>
<td>Em</td>
<td>Em</td>
<td>G</td>
<td>Em</td>
<td></td>
</tr>
</tbody>
</table>
### Case study

**What the final reporting looks like**

<table>
<thead>
<tr>
<th>Evaluation of stakeholder perceptions of implementation of Act</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent to which there is...</td>
<td>Poor</td>
</tr>
<tr>
<td>Administrative efficiency</td>
<td></td>
</tr>
<tr>
<td>Effective relationships</td>
<td></td>
</tr>
<tr>
<td>Good practice</td>
<td></td>
</tr>
<tr>
<td>Information, awareness and compliance</td>
<td></td>
</tr>
<tr>
<td>Overall rating</td>
<td></td>
</tr>
</tbody>
</table>

- **Green dot**: Excellent
- **Yellow dot**: Very good
- **Gray dot**: Good
We’ve come a long way….

Eoyang & Berkas 1999 provided a guide for:

- evaluating in uncertainty
- evaluator role in this context

Much holds true today
### Wide range of systems approaches & methods now used in evaluation

|----------------------------------------------|----------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| **Describing and analysing situations**      | ➢ Causal loop diagrams  
➢ System dynamics  
➢ Social network analysis  
➢ Outcome mapping                                                                 | ➢ Process monitoring of impacts  
➢ Strategic assumption surfacing and testing                                                                                       |
| **Changing and managing situations**         | ➢ Strategic areas assessment  
➢ The CDE Model  
➢ Assumption-based planning                                                               | ➢ Cynefin  
➢ Solution focus  
➢ Viable system model                                                                                                        |
| **Learning about situations**                | ➢ Cultural historical activity theory  
➢ Soft systems methodology  
➢ Critical systems heuristics                                                              | ➢ Scenario technique  
➢ Systemic questioning  
➢ Circular dialogues  
➢ Dialectical methods of inquiry                                                             |
BUT... still challenging terrain for evaluators

“It’s hard to ‘talk systems’ to most people without losing them immediately.”

Summary

- Systems thinking approaches of benefit - practical ideas:
  - Generic rubrics
  - Pattern spotters

- Approaches work for Government and NGO’s

- Eoyang & Berkas 1999 advice still holds.
Questions
You are welcome to cite this work.


I would like to acknowledge Glenda Eoyang’s support in providing input, advice and critique around the idea of rubrics as ‘simple rules’ and evaluating in uncertainty.
For further reference...


Davidson, E. J. (2012). Actionable evaluation basics: getting succinct answers to the most important questions [minibook]. Auckland: Real Evaluation Ltd.


