Making policy and making policy work with developmental evaluation

Conference or Workshop Item

How to cite:


© 2018 Martin Reynolds

Version: Accepted Manuscript

oro.open.ac.uk
Making policy and making policy work
With developmental evaluation

Martin Reynolds
Applied Systems Thinking in Practice (ASTiP) Group
The Open University, UK

Webinar 24th July 2018
Hosted by CECAN
Making policy and making policy work: overview of webinar presentation

• Developmental evaluation: complexity and practice
  – Policy evaluation and complexity
  – From external accountability towards internal responsibility
  – X8 principles (criteria for developmental evaluation)
  – X2 core criteria (complexity and systems thinking)

• Developmental evaluation: systems thinking in practice (STiP)
  – Policy evaluation and systems
  – X3 principles of systemic evaluation
  – X2 interplaying criteria (systemic and systematic)
  – STiP heuristic for enacting developmental evaluation

• Case story 1: evaluating postgraduate curriculum provision
• Case story 2: evaluating evaluation-in-practice

• Summary: making policy evaluation as public work
  – Journeying evaluation with systems thinking capabilities
  – Working principles for systems thinking in evaluation practice
1. ‘developmental’ evaluation is associated with any policy domain/situation (whereas ‘development’ evaluation is associated with specific policy domain/situation of international development)

2. ‘evaluation’: stakeholders in process of making value judgements on an evaluand (e.g. a situation, policy, or policy implementation)

3. ‘policy’ used as proxy to any intervention (including projects, programs, plans etc.) where the intention is to change or transform a situation of interest (e.g. primary health care support) associated with a domain of practice (e.g. health care provision)

4. ‘complexity’ relates to people; in turn relating to (1) situations being transformed, (2) stakeholding issues, (3) the actual policies devised to transform situations, and (4) the developmental path of ‘policy’ interventions.

Fig. 1 A mental model of policy and evaluation practice (devised by Reynolds)
(1980s…) Evaluation as ‘external accountability’: conventional evaluation practice

Derived from (conventional) 4-Step Logic Model of Evaluation (Scriven, 1980)…

1. Select criteria of merit and worth
2. Set performance standards (normative ‘ought’)

*Formative mode*

3. Collect data (analytical ‘is’) and compare with standards
4. Make value judgements

*Fig. 2* A mental model of evaluation as ‘external accountability’ (devised by Reynolds)
If policy intervention (design and implementation) is regarded as a (bounded) system
Evaluators are part of (internal to) the system


Attention to DE as (i) process dealing with complex situations, and (ii) requiring complexity tools and systems thinking
Systematic guidance derived from 20+ years experience; shift from ‘determining’ value towards ‘developing’ value.

1. Developmental purpose
2. Evaluation rigor
3. Utilization focus
4. Innovation niche
5. Complexity perspective
6. Systems thinking
7. Co-creation
8. Timely feedback
Two Key Principles for Developmental Evaluation: complexity perspective and systems thinking


1. Complexity perspective …x3 features
   - Conceptual ideas: emergence (self-organizing, attractors); nonlinear (small actions to large reactions… Butterfly effect); dynamic (interactive, volatile, changing); Getting to Maybe (uncertainty; unpredictable, uncontrollable; unanticipated consequences); co-evolutionary (interdependence between entities); adaptation (subject to continuous change)…
   - Context specific: ‘Complex’ situations are different from ‘simple’ and/or ‘complicated’ situations
   - Contingent (‘best fit’ for ‘innovation niche’…): developmental evaluation is not appropriate to all situations… only ‘complex’ evaluands (public health, cross-sector initiatives, social movements…); not ‘simple’ or ‘complicated situations

2. Systems thinking…x3 features/ orientations
   - Systems are ‘real’ (ontological devices): Systems are made up of sub-systems and function within larger systems; the whole is greater than the sum of the parts (forest vs trees);
   - Focus on interconnected relationships; Parts are interdependent such that a change in one part changes all parts…bias towards system dynamics (non-linear dynamics) tradition of systems thinking (cf. The Fifth Discipline, Peter Senge)
   - (core) conceptual tools: understanding evaluand as complex adaptive system (CAS); use of computerised agent-based modelling

(In contrast) systems thinking in practice (STiP) …. 
Systems thinking in practice (STiP)
...an alternative path for developmental evaluation

Three features / entities

1. Reality… (holistic domain of situations of interest) interdependent, non-linear etc. from which inevitably partial (incomplete) ‘factual’ judgements are made...

2. People … (pluralistic domain of evaluators and other stakeholders) with inevitably partial (biased) viewpoints expressed through individual value judgements (e.g. ‘viewing’ situations as ‘simple’)...

3. Systems …. (conceptual domain of constructs) used to simplify real world complexity, for purposes of:
   • Understanding evaluand (e.g. CAS)
   • Transforming evaluand (e.g. policy interventions)

Fig. 4 A mental model of three features of developmental evaluation
x3 principles of systemic evaluation

...based on interplay between facts, values, and boundaries

...and ideas of boundary critique and Systemic triangulation from Werner Ulrich (2003)


3. Accepting fallibility

(in making) Boundary judgements within the partial world of human activity

(ii) Partial in serving some stakeholder parties including practitioners - or interests - better than others

(i) Partial in representing only a section rather than the whole of the total universe of inter-relationships in any context that matters

2. Practicing empathy

(in making) Value judgements within ‘multiverse’ world of multiple stakeholders with multiple perspectives

1. Embracing humility

(in making) Factual judgements within the ‘universe’ of an interdependent and inter-related world of complicatedness, complexity & conflict

Fig. 5 An influence diagram illustrating three principles of systemic evaluation
Developmental evaluation heuristic

…through systems thinking in practice (STiP)

1. Situations of change: complicatedness, complex, and conflictual

2. People or stakeholders or practitioners … as agents of change

3. Ideas for example ‘systems’ and other conceptual
Tools as agency for change

(i) Understanding inter-relationships
…making/ developing factual judgements

(ii) Engaging with multiple perspectives
…making/ developing value judgements

(iii) Reflecting & appreciating limits on boundaries
of inter-relationships and perspectives …making/ developing boundary judgements

Reflecting on boundary judgements for purposes of
a) Systemic desirability
b) Cultural feasibility

Fig. 6 A mental model of systems thinking in practice (STiP) as a heuristic comprising three entities (situations, stakeholders, and systems) and three associated activities (uIR, eMP, and rBJ) …adapted from Reynolds and Howell (2010) Systems Approaches to Managing Change
A system for developmental evaluation
...practicing principles of systems thinking in practice

Fig 7 An activity model of a system to conduct developmental evaluation (adapted from a model of systemic inquiry: Checkland, 2002 and Ison, 2017)
Two case stories of developmental evaluation
...using systems thinking in practice (STiP)

Case story 1: Evaluating postgraduate curriculum provision


Case story 2: Evaluating evaluation-in-practice

Evaluating postgraduate curriculum provision

…using case study of postgraduate programme in systems thinking in practice (STiP)

Three developmental projects…

- **Project 1**: 2014-2016 (18 months)
  Enhancing systems thinking in practice in the workplace

- **Project 2**: 2016-2017 (12 months)
  Designing professional recognition for systems thinking in practice

- **Project 3**: 2018-2019 (18 months)
  Transforming postgraduate pedagogic praxis and workplace capabilities

---

**Fig. 8** Three nested systems of postgraduate curriculum development (adapted from Ison & Shelley, 2016 Fig. 1 p.589)
### Dimensions of STiP evaluation
- Understanding inter-relationships (uIR)
- Engaging multiple perspectives (eMP)
- Reflecting on boundary judgements (rBJ)

<table>
<thead>
<tr>
<th>Dimensions of STiP evaluation</th>
<th>Project 1 (18 months)</th>
<th>Project 2 (12 months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Enhancing systems thinking in practice (STiP) in the workplace</strong></td>
<td><strong>Designing professional recognition for systems thinking in practice (STiP)</strong></td>
</tr>
<tr>
<td></td>
<td><em>building capacity</em></td>
<td><em>...framing competencies</em></td>
</tr>
<tr>
<td>uIR: interviews +...</td>
<td><em>...postgrad students’ experiences in post-study workplace situations...x5 archetype</em></td>
<td><em>...workplace practices, professional practices &amp; Higher education provision...x20 (+) relevant competency framings (repository)</em></td>
</tr>
<tr>
<td>eMP: w/shops +...</td>
<td><em>...current students/ alumni/ employers...x5 archetype employer/alumnus relationships</em></td>
<td><em>...employers/ professional bodies associated with STiP...new model to support competency framing and capabilities</em></td>
</tr>
<tr>
<td>rBJ: reporting...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• systemically desirable</td>
<td>• <em>...to render ‘under the radar’ silent STiP practices/skills into more visibly acknowledged competencies</em></td>
<td>• <em>...to address tensions between systemic practices and systematic framing of competencies</em></td>
</tr>
<tr>
<td>• culturally feasible</td>
<td>• <em>...pluralist/ diverse culture of STiP practitioners</em></td>
<td>• *...changing role of Universities (corporate ‘good’ vs social ‘good’)</td>
</tr>
</tbody>
</table>
### Transforming curriculum praxis and capabilities

**…changing the way the game is played (Project 3)**

**Aim:** (Capabilities approach) shifting from developing ‘competencies’ based on learning outcomes (playing ‘the game’ better) towards enhancing ‘capabilities’ - creating innovative space for redefining occupational, professional, and social roles and practices amongst stakeholders in the workplace (changing the way ‘the game’ is played)?

<table>
<thead>
<tr>
<th>Dimensions of STiP</th>
<th>Project 3 (18 months)... to 2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>➢ Understanding inter-relationships (uIR)</td>
<td><strong>Transforming postgraduate pedagogic praxis and workplace capabilities: changing the way the game is played...developing capabilities</strong></td>
</tr>
<tr>
<td>➢ Engaging multiple perspectives (eMP)</td>
<td></td>
</tr>
<tr>
<td>➢ Reflecting on boundary judgements (rBJ)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>uIR: interviews +...</th>
<th>...understand systemic governance issues of curriculum design and implementation in relation to supporting part-time postgraduate study for enhancing workplace capabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>eMP: w/shops +...</td>
<td>...engage with progressing new Trailblazer Level 7 Apprenticeship standard involving multiple stakeholders including employers, professional bodies, other Higher Education providers</td>
</tr>
<tr>
<td>rBJ: reporting...</td>
<td>...draw on evaluative experiences of governance issues and the L7 apprenticeship for postgraduate STiP curriculum (re) design and implementation in 2020</td>
</tr>
</tbody>
</table>

- **systemically desirable**
- **culturally feasible**

”It is not about being the best at playing the game … but more about changing the way the game is played … while having fun in the process”

(Sports journalist, Guillem Balague, 16th April 2018. BBC)
1. understanding Inter-relationships (uIR)...

- Relational dynamics between evaluand (situations subject to evaluation), evaluators (making value judgements), and commissioners (decision makers responsible for overseeing evaluations)

- Use of ideas from (i) ‘systemic triangulation’ and boundary critique (Werner Ulrich, 2003) and (ii) ‘iron triangle’ (Ralph Pulitzer, 1919)

- Six activities of evaluation in practice revealed….

2. engaging Multiple perspectives (eMP)…over 20 years

- Research collaboration – since mid 1990s (evaluating participatory rural appraisal in Botswana), environmental planning, sustainability, conservation, public health, gender and equity, governance…

- Co-authoring – including systems thinking and complexity science in evaluation

- Conferencing (including round table forums), symposiums and webinars

- Workshop provision

- Consultancies

- Teaching

Evaluating evaluation-in-practice: two over-arching perspectives being sought…

a) (descriptive) What ‘is’ the situation?..*Evaluation-industrial complex (E-IC)* perspective

b) (normative) What ‘ought to be’ the situation?...*Evaluation-adaptive complex (E-AC)* perspective
3. Reflecting on boundary judgements (rBJ)…

<table>
<thead>
<tr>
<th>Actions evaluated ...associated with Evaluand (a) and (b)</th>
<th>Descriptive (actual) ‘is’ perspective Evaluation-industrial complex</th>
<th>Ideal (normative) ‘ought’ perspective Evaluation-adaptive complex (…systemically desirable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Audit Check: aspects of situation assessed</td>
<td>Situations systematically recognised as either simple, complicated (tame), or complex (wicked)</td>
<td>Situations systemically viewed as comprising all of (i) complicatedness (ii) complexity and (iii) conflict</td>
</tr>
<tr>
<td>b) Plan: terms of reference (ToR)</td>
<td>Purposive …fixed goals and targets as ascribed measures.</td>
<td>Purposeful - agile, flexible, adaptive measures</td>
</tr>
</tbody>
</table>

3. Reflecting on boundary judgements (rBJ)…

<table>
<thead>
<tr>
<th>Actions evaluated ...associated with Evaluators (c) and (d)</th>
<th>Descriptive (actual) ‘is’ perspective Evaluation-industrial complex</th>
<th>Ideal (normative) ‘ought’ perspective Evaluation-adaptive complex (…systemically desirable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>c) Evaluation summative: criteria used (measures)</td>
<td>Evaluator ‘external’ to evaluand:</td>
<td>Evaluator part of evaluand;</td>
</tr>
<tr>
<td></td>
<td>Focus more on efficacy/worth and efficiency/merit ('outputs' and doing things right) as pre-set criteria, rather than effectiveness/worthiness/significance ('outcomes and impacts' and doing the right thing).</td>
<td>Able to continually juggle between criteria of efficacy, efficiency, and effectiveness (including ethical and political notions of equity and sustainability). Includes intrinsic ‘personal’ values/principles.</td>
</tr>
<tr>
<td></td>
<td>Stress on impartial (quasi) positivist epistemology.</td>
<td>Stress on partiality of constructivist epistemology.</td>
</tr>
<tr>
<td>d) Evaluation formative: tools used, values developed</td>
<td>Evaluator ‘external’ to evaluand: primary specialist ‘scientific’ role of measuring value as part of either repeated use of same fixed tools - 'best practice', or seeking ever growing ‘toolbox’ adopting new tools as ‘best fit’ for purpose ('horses for courses').</td>
<td>Evaluator part of evaluand; more generalist role as an agile ‘bricoleur’ a crafts person formatively developing value - instrumental (utility), intrinsic (rights-based), and personal (justice) - of stakeholders adapting existing tools for purpose.</td>
</tr>
<tr>
<td></td>
<td>‘power-over’ = dominant power relation attribute...setting criteria for later summative evaluation</td>
<td>‘power-to’ and ‘power-with’ and ‘power-within’ (empowerment) = dominant attributes</td>
</tr>
</tbody>
</table>
3. Reflecting on boundary judgements (rBJ)...

<table>
<thead>
<tr>
<th>Actions evaluated</th>
<th>Descriptive (actual) ‘is’ perspective Evaluation-industrial complex</th>
<th>Ideal (normative) ‘ought’ perspective Evaluation-adaptive complex (…systemically desirable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>e) Commissioning guarantors of rigour: assurances, trustworthiness, and responsibility</td>
<td>Guarantor of <em>truthfulness</em> through objective ‘evidence’ … objective reliable and replicable use of tools through data triangulation (multi methods)</td>
<td>3 sets of co-guarantor attributes, for <em>truthfulness</em> through appropriate <em>deliberation</em> – (i) reliability (multidisciplinary), (ii) resonance (interdisciplinary complementarity or—communicable with other groups/cultures etc,) and (iii) relevance (transdisciplinary dialogue with wider social and ecological concerns)</td>
</tr>
<tr>
<td></td>
<td>Responsibility limited towards accountability to decision makers</td>
<td>Responsibility involves attributes of caring as well as accountability</td>
</tr>
<tr>
<td>f) Learning developed</td>
<td>Mostly single-loop (is the intervention being done right?) and occasional double-loop learning (is it doing the right thing?) expressed, but generally less reflective of power relations circumscribing the intervention and/or circumscribing the use of tools for evaluating the intervention.</td>
<td>Single-loop, double-loop and triple-loop learning are all evident. Intervention regarded as political, with awareness of, and adaptive address to, power relations affecting the intervention and being effected through the intervention.</td>
</tr>
<tr>
<td></td>
<td>Evaluation regarded as apolitical</td>
<td>Evaluation regarded as ‘political’ act</td>
</tr>
</tbody>
</table>
Opportunities:

- Complexity of interventions (policy/programmes/projects...involving both design & implementation) acknowledged and more appreciated in most sectors
- Evaluation increasingly regarded as integral to any intervention
- Evident need for more simple heuristics to work appropriately with complexity (Matthew Taylor RSA)
- Importance of 'language' in conversations between evaluators and policy making communities of (Siobhan Campbell)
- Evaluations and humility... claims are more circumspect
- 'developing value’ …Principles of evaluand i.e. Principle-focused evaluation (Patton, 2018)
- Evaluation as political (deliberative evaluation…) cf. Thomas Schwandt (from ‘what should be done?’ towards ‘what should we do?’); ISE4GEMs (Lewis and Stephens, 2017)
- Push-back against 'expertocracy’ (but risks of neoliberal populism...)

Challenges:

- Evaluation as ‘external accountability’ still dominant
- Turbulent times: post-truth (evidence based fatigue) or rather post-trust?
- ‘Think like a system’ (?) ‘Act like an entrepreneur’ … Matthew Taylor
- Prevalence of ‘contingency thinking’ (simple/ complicated or complex)
- Methodological/ method fetishism (social sciences…); burgeoning ‘tool box’.. Empathy with users(?)
- Complexity ‘tools’ as silver bullets...
- Prevalence of dualisms – either ‘facts’ or ‘values’ …evidence or meaning...
Summary 1: Journeying evaluation with systems capabilities approach to evaluation praxis

Heinz von Foerster: ethical action is to ‘act always so as to increase the number of choices’

…and increase ability to appropriately choose and develop value in the process

A core capability for systems thinking in practice is praxis… avoiding dualisms from dualities
Journeying evaluation with systems
...capability to appropriately be systemic and systematic

Systems thinking as iteration between:

1. Systemic... understanding real world (of complicatedness, complexity, and conflict)... theory

2. Systematic... engaging real world (e.g. listening to different perspectives)... practice

Not 'either/or' but 'both/and' (ying and yang)
Journeying evaluation with systems
...capability towards ontological and epistemological use of systems

Iterating between two devices for thinking about ‘systems’

1. Ontological device: systems to be evaluated
   ‘the’ health system
   ‘the’ legal system etc.
   (…regarded as complex adaptive systems)
   
   purposive engagement (e.g. external accountability)

2. Epistemological device: situations to be evaluated using systems as a learning device
   
   purposeful engagement using systems design

Not ‘either/or’ but ‘both/and’ (ying and yang)
Journeying evaluation with systems
Developing praxis capability with systems thinking in practice

‘juggling’ with (2 balls of…) dualities

- value judgements and factual judgements
- epistemological drive (knowing) and ontological drive (knowns/unknowns)
- systematic and systemic
- perspective/perception and inter-relationships, perspectives, boundaries
- engaging multiple perspectives (eMP) and understanding inter-relationships (uIR)
- practice and knowledge
- action and research
- meaning and evidence
- humanities and sciences
- trust and truth
- empathy and humility

All mediated through reflection on (3rd ball of…) boundary judgements
= systems praxis

+ having fun in the process

Systemic triangulation: adapted from Werner Ulrich (2003)
X3 ethical principles of systems thinking in evaluation practice

1. Embrace humility (inter-relationships)
2. Practice empathy (perspectives)
3. Accept fallibility (boundary judgements)

X6 ‘operating principles’ of systems thinking in evaluation practice

1. (audit) start systemically (complications, complexities, and conflict)
2. (plan) keep objectives flexible in time (cf. adaptive action.. ‘what/ so what/ now what’.. Glenda Eoyang)
3. (evaluation summative) attend to ethical criteria of wellbeing… who might be the victims?
4. (evaluation formative) attend to power relations (privileging power-to and power-with)
5. (commissioning) provide robustness/ rigour without rigor-mortis (trapped in one co-guarantor of ‘objectivity’ at expense of other co-guarantors
6. (learning) generate learning that questions ethics (doing the right thing) as well as politics (power and knowledge…who determines what’s right?)
Contact details and resources

Martin Reynolds
Qualifications Lead (PG systems thinking in practice)
School of Engineering and Innovation
The Open University
Walton Hall
Milton Keynes MK7 6AA
Email: martin.reynolds@open.ac.uk
Website Publications Resources