Evaluating Systems Change: An Inquiry Framework
1. Context

Types of Evaluation | Principles Versus Recipes | Decision-Making Pie
More and more resources

DESCRIPTING. Systems Concepts in Action by Williams & Hummelbrunner.


PRINCIPLES. Principles for Evaluating Complexity. Preskill & Gopal et al.

METHODS. Center for Evaluation Innovation, by Coffman & Beer

OVERALL ORIENTATION. Developmental Evaluation Exemplars by Patton, McKegg, Wehipeihana

CASE STUDIES. Regularly added to Better Evaluation website and journals of evaluation.
.... and some very interesting domain specific results frameworks emerging

- Health Systems
- Food Security
- Workforce Development
- Climate Change
- Child Welfare System
- Education System
- Justice System
- Housing
- And many more
... but shortage of purpose, framing and principles.
### Evaluation Purpose

#### 6 Major Types of Evaluation

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Priority Questions</th>
<th>Common Methods</th>
<th>Key Factors Affecting Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developmental</td>
<td>What is happening in the context in which the program operates?</td>
<td>Emergent evaluation, Complexity frameworks</td>
<td>Innovators interest in using data to inform their decisions</td>
</tr>
<tr>
<td></td>
<td>What can we control and not control?</td>
<td>Ex ante evaluation</td>
<td>Tolerance for ambiguity and uncertainty</td>
</tr>
<tr>
<td></td>
<td>What is developing?</td>
<td></td>
<td>Balancing quality and speed of feedback</td>
</tr>
<tr>
<td></td>
<td>What are the effects of the program?</td>
<td></td>
<td>Integrate and synthesize multiple and conflicting sources of data</td>
</tr>
<tr>
<td></td>
<td>What are the options for the next iteration of the program?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formative</td>
<td>What is working and what is not?</td>
<td>Quality enhancement, Learning reviews, Reflective practice</td>
<td>Creating a learning climate, openness to feedback and change</td>
</tr>
<tr>
<td></td>
<td>What are the strengths and weaknesses of the program?</td>
<td>Participant feedback, Appreciative Inquiry</td>
<td>Trust</td>
</tr>
<tr>
<td></td>
<td>What works for whom, in what context, and why?</td>
<td></td>
<td>Evaluators skills in facilitated learning</td>
</tr>
<tr>
<td></td>
<td>How can results be increased and costs reduced?</td>
<td></td>
<td>Relevance of findings</td>
</tr>
<tr>
<td>Summative</td>
<td>To what extent does the program achieve its goals?</td>
<td>Impact Evaluation, Cost-Benefit Analysis, Randomized Controlled Trials</td>
<td>Independence and credibility of the evaluator</td>
</tr>
<tr>
<td></td>
<td>Does the program have value for money?</td>
<td></td>
<td>Rigor of the design: validity and generalizability</td>
</tr>
<tr>
<td></td>
<td>How do its outcomes and costs compare with other options?</td>
<td></td>
<td>Significance of the findings to decision-makers</td>
</tr>
<tr>
<td></td>
<td>To what extent can outcomes be attributed to the interventions?</td>
<td></td>
<td>Timeliness</td>
</tr>
<tr>
<td></td>
<td>Should this program be sustained, wound down, replicated or scaled?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accountability</td>
<td>Are funds being used for intended purposes?</td>
<td>Program audits and inspections, Performance measurement and monitoring</td>
<td>Government and funder mandated reporting</td>
</tr>
<tr>
<td></td>
<td>Are goals and targets being met?</td>
<td>Accreditation and licensing, End-of-project lists, Scorecards</td>
<td>Validated indicators</td>
</tr>
<tr>
<td></td>
<td>Are problems being handled?</td>
<td></td>
<td>Integrity and credibility of the system and those reporting</td>
</tr>
<tr>
<td></td>
<td>Is implementation following the approved plan?</td>
<td></td>
<td>Balance and consistency of reporting</td>
</tr>
<tr>
<td></td>
<td>Are quality-control mechanisms in place and being used?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monitoring</td>
<td>Are inputs and processes flowing smoothly?</td>
<td>Management information systems, Quality control systems, Continuous quality improvement, Six Sigma, routine reporting, performance indicators</td>
<td>Timeliness, regularity, relevance, and consistency of reporting</td>
</tr>
<tr>
<td></td>
<td>Are participation and dropout rates?</td>
<td></td>
<td>Incentives to input data at field level and to use it at management level</td>
</tr>
<tr>
<td></td>
<td>Are these changing?</td>
<td></td>
<td>Capacity and resources to manage system</td>
</tr>
<tr>
<td></td>
<td>Are outputs being produced as anticipated and scheduled?</td>
<td></td>
<td>Appropriate links to accountability system</td>
</tr>
<tr>
<td></td>
<td>What bottlenecks are emerging?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>What variations occur across sub-groups?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge Generation</td>
<td>What are the general patterns and principles of effectiveness in programs, projects and sites?</td>
<td>Cluster evaluation, Meta-analyses, Synthesis evaluation</td>
<td>Quality and comparability of source used</td>
</tr>
<tr>
<td></td>
<td>What lessons being learned?</td>
<td>Lessons Learned, Effective practice studies</td>
<td>Quality of synthesis</td>
</tr>
<tr>
<td></td>
<td>What principles can be extracted across results to inform practice?</td>
<td></td>
<td>Capacity of users to extrapolate</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rigour of triangulation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Identifying principles that can inform practice</td>
</tr>
</tbody>
</table>

Principles

**WHAT WE KNOW SO FAR ABOUT**

**Sets of Principles for Evaluating Systems Change Efforts**

The idea of using principles to guide evaluation is not new. Evaluators learned a long time ago that it was difficult to describe the “best” design or methods: an evaluation to determine if a Bangladesh micro-lending program is replicable in a New York City neighbourhood will differ from one conducted by a health ministry in the United Kingdom, that uses rapid improvement cycles to reduce wait times in hospital emergency rooms. There is an endless variety in issues, interventions, user questions, available resources, evaluator skill, etc. It is impossible to arrive at a single design that can address all of them.

As a result, evaluators have increasingly focused on developing universal principles that identify key considerations to keep in mind when carrying out an assessment. It is then up to evaluators and innovators to figure out how to apply these principles, case by case. Take, for example, the simple principle: “We cannot evaluate everything we want to assess, so we must focus our attention on what is important.” How that looks in practice depends entirely on the unique context of an initiative.

Evaluation principles are now popping up everywhere. The Hewlett Foundation created its own set to help design internal (and contracted) evaluations of its grant-making programs. (Eknayan & Lindheim 2007; In Canada, the Collaborative Opportunities to Value Evaluation group has developed principles to guide collaborative approaches to assessment (Shulha et al. 2016). Kettenman (1994) has put together a set of principles to guide an empowerment model of evaluation. Once you start looking for evaluation principles, you can see them everywhere.

Yet, despite this mainstreaming of principles-based approaches to evaluation, there has been relatively little work on creating principles to guide the evaluation of systems change efforts. The reason is simple: evaluation practitioners are a few steps behind the social innovations’ rapid and widespread adoption of systems change strategies. It’s a bit like “catch up.”

Thankfully, there has been progress. The table below describes three sets of principles developed by leaders in evaluation in North America (and beyond). They are excellent. I refer to them regularly. However, I also draw on my own set which I created and refined 2001-2011 while directing the Vibrant Communities Initiative, a network of 15 urban collaborators seeking to reduce poverty. Much of its work concerned changing

<table>
<thead>
<tr>
<th>Principles</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developmental Evaluation Examples: Principles in Practice (Patton, Mckegg, &amp; Warkentin 2015)</td>
<td>12 principles focused on systems change, with an extra focus on human services systems.</td>
</tr>
<tr>
<td>Evaluating Complexity: Propositions for Practice (Freire, Gopin, Merx, Cook 2014)</td>
<td>Propositions that are embedded in a complexity world view of sluggish social, economic, and environmental challenges.</td>
</tr>
<tr>
<td>Framework for Evaluating Systems Change (Coffman 2007)</td>
<td>If principles to guide complexity-based, developmental evaluation, one of which includes “systems” lenses.</td>
</tr>
</tbody>
</table>
Key Framing Points

1. Why change systems?
2. What is a system?
3. What do results in a system change efforts look like?
4. What are the implications of disruptive events?
5. So what are the implications and questions for you work?
1. Why change systems?

Achieving impact at scale.
The Case for Systems Change

*Programmatic interventions help people beat the odds.*

*Systemic interventions can help change their odds.*

Karen Pittman
CEO, Forum on Youth Investment
Keynote at Tamarack Community Change Institute
Vancouver, 2015
The Case for Systems Change

Systems change is about shifting the conditions that hold a problem in place.

Social Innovation Generation Canada
2. What is a system?

Boundaries + Actors + Relationships & More
What does system change mean?

I was asked to work with innovators in the national health program of a smaller country. When I started working with the group, they said, ‘We aim to shift the health system.’ After listening for a few hours, I said, ‘Honestly, I have no idea what you are doing, or what you are trying to achieve … and I haven’t a clue how to measure it. I don’t understand what it means to “shift the health system.”’ And they looked at each other and burst out laughing and said, ‘We have no idea either.’

Donna Podems
Experienced Evaluator in South Africa
Key Features

- **Boundaries** (i.e. domain, geographic, scale, etc.)

- **Agents/Actors** (e.g. diverse organizations, people, networks, etc.)

- **Relationships** (e.g. tight-loose, hierarchical-distributed, power, legitimacy, etc.)

Complex Dynamics

- **Overlapping/Open** – one system is entangled up and influenced by other systems.

- **Emergent behaviour** – system behaviour is the sum of all actors’ interactions and actions.

- **Non-Linear dynamics** – inputs and outputs/outcomes not clearly related.

- **Varying levels of predictability** – a splatter and ripple of effects.

- **Ever-evolving** – does not stand still.
3. What do results look like?

Systems Change | Mission Outcomes | Strategic Learning
Multiple Results

A framework that provides social innovators, evaluators and funders to consider three types of results they need to inform their ongoing efforts to change systems underlying tough challenges.

**SYSTEMS CHANGE**
The extent to which efforts change the systems underlying complex issues.

1. Changes in drivers of system behaviors
2. Changes in behaviors of system actors
3. Changes in overall system behavior(s)

**MISSION OUTCOMES**
The extent to which our efforts help to make lives better.

1. Outcomes for individuals
2. Outcomes for targeted geography/groups
3. Outcomes for populations

**STRATEGIC LEARNING**
The extent to which efforts uncover insights key to future progress.

1. Learning about what we are doing
2. Learning about how we are thinking
3. Learning about how we are being
Multiple Results

A framework that provides social innovators, evaluators and funders to consider three types of results they need to inform their ongoing efforts to change systems underlying tough challenges.

**SYSTEMS CHANGE**
The extent to which efforts change the systems underlying complex issues.
1. Changes in drivers of system behaviors
2. Changes in behaviors of system actors
3. Changes in overall system behavior(s)

**MISSION OUTCOMES**
The extent to which our efforts help to make lives better.
1. Outcomes for individuals
2. Outcomes for targeted geography/groups
3. Outcomes for populations

**STRATEGIC LEARNING**
The extent to which efforts uncover insights key to future progress.
1. Learning about what we are doing
2. Learning about how we are thinking
3. Learning about how we are being
Three Types of Outcomes

#1: Changing Drivers

#2: Changing Actor Behaviors

#3: Changing Overall System Behavior
Example: Housing First
Three Types of Outcomes

#1: Changing Drivers

#2 Changing Actor Behaviors

#3 Changing Overall System Behavior
Case Study: Medicine Hat

Medicine Hat Homelessness Could Reach Its End This Year

The Huffington Post Alberta | By Jesse Ferreras  
Posted: 05/14/2015 10:20 am EDT  |  Updated: 05/15/2015 11:59 am EDT

Medicine Hat is about to end homelessness – but then what?

ANNALISE KLINGBEIL, CALGARY HERALD

More from Annalise Klingbeil, Calgary Herald

Published on: April 15, 2015 | Last Updated: May 12, 2015 8:47 PM MDT
Multiple Results

A framework that provides social innovators, evaluators and funders to consider three types of results they need to inform their ongoing efforts to change systems underlying tough challenges.

**SYSTEMS CHANGE**
The extent to which efforts change the systems underlying complex issues.
1. Changes in drivers of system behaviors
2. Changes in behaviors of system actors
3. Changes in overall system behavior(s)

**MISSION OUTCOMES**
The extent to which our efforts help to make lives better.
1. Outcomes for individuals
2. Outcomes for targeted geography/groups
3. Outcomes for populations

**STRATEGIC LEARNING**
The extent to which efforts uncover insights key to future progress.
1. Learning about what we are doing
2. Learning about how we are thinking
3. Learning about how we are being
Case Study #2: Payday Lending Reform in Alberta

A collaborative response led by Momentum (Calgary), one of Canada’s most sophisticated Community Economic Development organizations, in partnership with many, many others.
Media Strategy
City Task Force on Financial Empowerment
Changes to City Business & Zoning By-Laws
Progressive Financial Institution Engagement
Provincial Act on Predatory Lending developed & passed.
Dramatic changes in Payday lending activities.
Advocacy begins with Federal Government Bill on Usury Rates
Neighborhood Organizing
Media Strategy
Champion Development

Navigating the LANDSCAPE
Shifting AWARENESS & CULTURE
Nudging SYSTEMS
Supporting NICHE INNOVATIONS
Strengthening CAPACITY & AGENCY

Cash Crunch Loans

MISSION IMPACT
Multiple Results
A framework that provides social innovators, evaluators and funders to consider three types of results they need to inform their ongoing efforts to change systems underlying tough challenges.

SYSTEMS CHANGE
The extent to which efforts change the systems underlying complex issues.
1. Changes in drivers of system behaviors
2. Changes in behaviors of system actors
3. Changes in overall system behavior(s)

MISSION OUTCOMES
The extent to which our efforts help to make lives better.
1. Outcomes for individuals
2. Outcomes for targeted geography/groups
3. Outcomes for populations

STRATEGIC LEARNING
The extent to which efforts uncover insights key to future progress.
1. Learning about what we are doing
2. Learning about how we are thinking
3. Learning about how we are being

POPULATION
• Participants in pilot programs obtain lower cost loans
• Borrowers report liking installment model

TARGETED
• The # of payday lending stores on International Avenue drops from 11 to 3.
• Unknown impact for different demographic groups.

INDIVIDUALS
• Participants in pilot programs obtain lower cost loans
• Borrowers report liking installment model

INDIVIDUALS
• Participants in pilot programs obtain lower cost loans
• Borrowers report liking installment model

POPULATION
• The # of payday stores in the Province dropped from 230 to 165, and volume of lending from $500 mil to $285 mln, since the Payday lending bill was passed.

POPULATION
• The # of payday stores in the Province dropped from 230 to 165, and volume of lending from $500 mil to $285 mln, since the Payday lending bill was passed.

POPULATION
• The # of payday stores in the Province dropped from 230 to 165, and volume of lending from $500 mil to $285 mln, since the Payday lending bill was passed.

POPULATION
• The # of payday stores in the Province dropped from 230 to 165, and volume of lending from $500 mil to $285 mln, since the Payday lending bill was passed.

POPULATION
• The # of payday stores in the Province dropped from 230 to 165, and volume of lending from $500 mil to $285 mln, since the Payday lending bill was passed.
Multiple Results

A framework that provides social innovators, evaluators and funders to consider three types of results they need to inform their ongoing efforts to change systems underlying tough challenges.

**SYSTEMS CHANGE**
The extent to which efforts change the systems underlying complex issues.
1. Changes in drivers of system behaviors
2. Changes in behaviors of system actors
3. Changes in overall system behavior(s)

**MISSION OUTCOMES**
The extent to which our efforts help to make lives better.
1. Outcomes for individuals
2. Outcomes for targeted geography/groups
3. Outcomes for populations

**STRATEGIC LEARNING**
The extent to which efforts uncover insights key to future progress.
1. Learning about what we are doing
2. Learning about how we are thinking
3. Learning about how we are being
Case Study: Edmonton’s Shift Lab - Human Centered Design to Address Racism & Poverty in the Domain of Housing

THE APPROACH

1. EMPATHY
   - Stories Ethnographic Research
   - Sense Making
   - System Mapping

2. DEFINE
   - Making sense of needs and insights from stories “How Might We” Questions

3. IDEATE
   - Brainstorming
   - Getting ideas from other fields
   - Co-designing with community
   - Building on ideas of others

4. PROTOTYPE
   - Checking the prototypes with community/wi user groups the prototypes are for

5. TEST
   - Choosing ideas that could meet needs
   - Making prototypes of what a service, policy innovation could look like

HUMAN CENTRED LAB PROCESS

THE PROTOTYPES

- Mobile Legal Aid
- Diversity Certification Program for Housing Suppliers
- Journey to YIMBY (yes in my back yard)
• 11 key insights about the strengths & limitations of workshop design, facilitation team management, pacing of sessions, etc.

• 8 insights about the nature of racism, poverty, and housing in Edmonton: e.g., racism can be internalized, interpersonal, & systemic
• 3 insights about Lab methodology: e.g., relationships between innovation and social justice models of change

• We struggle to have deeper, more meaningful & productive conversations about racism & how that limits our ability to move forward

Mission Critical Triple Loop Learning

We mostly stayed here and therefore did not go deep enough with our thinking or interventions.

4 Types of Conversations

- **Generative Dialogue**
  - presencing, flow
  - time: slowing down
  - space: boundarystake
  - listening from one’s Future Self
  - rule-generating

- **Reflexive Dialogue**
  - inquiry
  - I can change my view
  - empathic listening (from within the other self)
  - other – you
  - rule-reflecting

- **Talking Nice**
  - downloading
  - polite, cautious
  - listening – projecting
  - rule-revealing

- **Talking Tough**
  - debate, clash
  - I am my own point of view
  - listening – reloading
  - other = target
  - rule-revealing
Shift Lab 2.0: Adaptation

• supporting the evolution of the original three prototypes

• includes a strong emphasis on developing capacity for tough and important conversations about racism in Edmonton

• seeks to prototype behavioral change interventions to encourage the ‘sleepy middle’ of Edmonton to see, understand and respond to racism
Some Resources on Strategic Learning

• https://www.fsg.org/areas-of-focus/strategic-learning-evaluation
Strategic Learning Resources

• The Foundations Strategy Group - [https://www.fsg.org/areas-of-focus/strategic-learning-evaluation](https://www.fsg.org/areas-of-focus/strategic-learning-evaluation) – a good set of basic resources.

• The Center for Evaluation Innovation - [https://www.evaluationinnovation.org/topics/](https://www.evaluationinnovation.org/topics/) - veteran evaluators with some case studies and techniques.

• Philanthropic Review - [https://johnsoncenter.org/resources/thefoundationreview](https://johnsoncenter.org/resources/thefoundationreview) - includes good examples of strategic learning in action.
Multiple Results
A framework that provides social innovators, evaluators and funders to consider three types of results they need to inform their ongoing efforts to change systems underlying tough challenges.

What are your initial impressions of the framework?

<table>
<thead>
<tr>
<th>SYSTEMS CHANGE</th>
<th>MISSION OUTCOMES</th>
<th>STRATEGIC LEARNING</th>
</tr>
</thead>
<tbody>
<tr>
<td>The extent to which efforts change the systems underlying complex issues.</td>
<td>The extent to which our efforts help to make lives better.</td>
<td>The extent to which efforts uncover insights key to future progress.</td>
</tr>
<tr>
<td>1. Changes in drivers of system behaviors</td>
<td>1. Outcomes for individuals</td>
<td>1. Learning about what we are doing</td>
</tr>
<tr>
<td>2. Changes in behaviors of system actors</td>
<td>2. Outcomes for targeted geography/groups</td>
<td>2. Learning about how we are thinking</td>
</tr>
<tr>
<td>3. Changes in overall system behavior(s)</td>
<td>3. Outcomes for populations</td>
<td>3. Learning about how we are being</td>
</tr>
</tbody>
</table>
5. What are the implications of disruptive events?

Uncertain Futures
Disruptions & Uncertainty
Three Horizons Revisited

- This adapted version of the Three Horizons Framework illustrates that the various horizons for developing innovative responses to complex issues differs changes from stable to crisis contexts.

## Three Horizons: Stable versus Crisis Contexts

<table>
<thead>
<tr>
<th>Horizon 1: Incremental Innovation</th>
<th>Stable Contexts</th>
<th>Crisis Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>• **Aim to develop ideas, actions and practices that <strong>improve - but don’t disrupt</strong> - the current system</td>
<td><strong>Horizon 1: Crisis Management</strong></td>
<td>• Identify and respond to and mitigate the ripple effects of the crisis, with an emphasis on mitigating the (triage) while adaptively managing critical routine operations</td>
</tr>
<tr>
<td>• Level of effort is routine; results are relatively predictable, low risk, immediate, yet modest</td>
<td></td>
<td>• Level of effort is considerable: results are real time &amp; emergent, range from proactive to reactive, unpredictable and often high stakes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Horizon 2: Disruptive Innovation</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Seek to disrupt the systems that hold complex challenges in place through bolder experiments, deliberate pushes to change policies, power structures, public awareness</td>
<td><strong>Horizon 2: Post Disruption Recovery</strong></td>
<td>• Track, develop and understand new ideas and practices developed during the crisis that can be employed more readily – and perhaps even help prevent – a future iteration of the crisis</td>
</tr>
<tr>
<td>• Level of effort is far greater, uncertain and involve tension and risk; results require longer to emerge, are unpredictable, but may be significant</td>
<td></td>
<td>• Level of effort is modest in the early days of the crisis but gradually increases as the emergency softens and there is more time and receptivity to think about post-crisis recovery; results appear post recovery</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Horizon 3: Transformative Innovation</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Generate and promote alternative visions of a future world with new ideas learning rich exemplars and widespread engagement.</td>
<td><strong>Horizon 3: Post Recovery Transformation</strong></td>
<td>• Generate and amplify older and new visions and ideas for significant change, build and engage networks, seek and exploit newly opened windows of opportunity</td>
</tr>
<tr>
<td>• Level of effort and results are typically diffuse, long term and difficult to track; potential pay-off is unpredictable and ranges from nothing to the transformative</td>
<td></td>
<td>• Level of efforts and results are mix of fast and slow, diffuse and concentrated, risky and not, yet always unpredictable; results of adopted ideas can be transformative</td>
</tr>
</tbody>
</table>
What ideas do you want to inform your post-Pandemic world?

• “Only a crisis - actual or perceived - produces real change. When that crisis occurs, the actions that are taken depend on the ideas that are lying around. That, I believe, is our basic function: to develop alternatives to existing policies, to keep them alive and available until the politically impossible becomes the politically inevitable.”
  
  — Milton Friedman
1. Why change systems?
2. What is a system?
3. What do results in a system change efforts look like?
4. What are the implications of disruptive events?
5. So what are the implications and questions for you work?
Group Discussion

So what insights emerge for you regarding your own efforts to track and communicate your efforts change systems?

So what new questions emerge?

**SYSTEMS CHANGE**
The extent to which efforts change the systems underlying complex issues.
1. Changes in drivers of system behaviors
2. Changes in behaviors of system actors
3. Changes in overall system behavior(s)

**MISSION OUTCOMES**
The extent to which our efforts help to make lives better.
1. Outcomes for individuals
2. Outcomes for targeted geography/groups
3. Outcomes for populations

**STRATEGIC LEARNING**
The extent to which efforts uncover insights key to future progress.
1. Learning about what we are doing
2. Learning about how we are thinking
3. Learning about how we are being