Getting to Maybe: Evaluation, Systems Thinking, and Complexity Science

Evaluation History

Setting the Context: Context Matters

So, In the beginning…


• 1st edition, 1978
• 2nd edition, 1986
• 3rd edition, 1997

New Direction # 1

International and cross-cultural expansion of evaluation: globalization and diversity

New Direction # 2

From Studies to Streams
Evaluation Trends:
20 years ago
- One study for one user
- Modest databases
- Long time frames for studies
- Presumption of direct use
- Long reports
- Generally single method
- Dissemination was the written word

10 years ago
- One study for multiple users.
- The age of stakeholders
- Larger databases with computer support
- More client focus
- Not just decision use, but also conceptual use
- Multiple teams producing information
- Quantitative/qualitative wars come to an end: valuing multiple methods

And now...
- We are moving from discrete studies to information streams
- Systems not individual evaluators produce evaluative knowledge
- Evaluative streams are multiple – integrating information from different sources

And now...
- Data collection & reporting at multiple levels by multiple-stakeholders
- Databases are continuous and virtual
- Time frames are immediate
- Analysis is continuous
- Virtual analysis of trends and conditions
- Visual displays instead of narratives

And now
- Partnerships are dominant in collecting, analyzing and sharing evaluative knowledge
- Internet is the new information glue
- Increased transparency of evaluative knowledge
- Emphasis on continuous organizational adaptation and improvement

New Direction # 3
Proliferation of evaluation models, theories, options, and methods, and approaches
Original Primary Options

Formative and Summative Evaluation
(Mid-term and End-of-Project Reviews)

Blandin Community Leadership Program
Developmental Evaluation

Evidence-based Practice
Evaluation grew up in the “projects” testing models under a theory of change that pilot testing would lead to proven models that could be disseminated and taken to scale:

The search for best practices and evidenced-based practices

Fundamental Issue: How the World Is Changed
Top-down dissemination of “proven models” versus Bottoms-up adaptive management

Conditions that challenge evaluation
• High innovation
• Development
• High uncertainty
• Dynamic
• Emergent
• Systems change

Challenge:
Matching the evaluation process and design to the nature of the situation:
Contingency-based Evaluation
New Direction # 4
Broader understanding and conceptualization of evaluation use

Conceptualizing Use
• Utilization-focused evaluation now includes knowledge management, organizational learning, and facilitating change. The focus is as much on institutional uses of knowledge as on individual users

Re-conceptualizing Use
• Use is a process not an event
• Use involves an interaction not just a report
• Use involves training for use not just delivery of results
• Use begins at the beginning not at the end

New Direction # 5
Increased up-front role for evaluation & evaluators in intervention design:
Logic modeling & Theory of change work

Some premises:
• Evaluation is part of initial program design, including conceptualizing the theory of change
• Evaluator’s role is to help users clarify their purpose, hoped-for results, and change model.
• Evaluators can/should offer conceptual and methodological options.
• Evaluators can help by questioning assumptions.
• Evaluators can play a key role in facilitating evaluative thinking all along the way.
• Designs can be emergent and flexible.

New Direction # 6
Beyond linear logic models:
Systems Thinking and Complexity Science
Three ways of conceptualizing and mapping theories of change

- Linear Newtonian causality
- Interdependent systems relationships
- Complex nonlinear dynamics

Linear Logic Model

INPUTS (people, materials) → ACTIVITIES (processes) → OUTPUTS → OUTCOMES → CHANGES IN PEOPLES LIVES → IMPACTS → CHANGES IN COMMUNITIES

Causal Diagram of Beckhard’s Change Formula


Dissatisfaction with the Status Quo → Resistance to Change → Compelling Vision

Believability → First Steps

Feedback Systems Logic Model

Inputs → Key Processes & Factors → Outputs/Outcomes → Customer/Organization Interaction

Customer/Organization Interaction → Feedback to Key Processes/Factors → Inputs

Division for Head Disease and Stroke Prevention/Early Warning Logic Model

Inputs → Processes → Outputs/Outcomes → Feedback Loop
Step Two: THEN A MIRACLE OCCURS
“I think you should be more explicit here in step two.”

Systems

- Parts are interdependent such that a change in one part changes all parts
- The whole is greater than the sum of the parts
- Focus on interconnected relationships
- Systems are made up of sub-systems and function within larger systems

Bob Williams and Iraj Imam
AEA Monograph,
EdgePress/AEA Point Reyes CA.

Understanding the Elephant from a Systems Perspective
The relationship between what goes in and what comes out

What conceptual framework informs front-end evaluation work?

Teen Pregnancy Program Example

Logic Model for Pregnant Teens Program
1. Program reaches out to pregnant teens
2. Pregnant teens enter and attend the program (participation)
3. Teens learn prenatal nutrition and self-care (increased knowledge)
4. Teens develop commitment to take care of themselves and their babies (attitude change)
5. Teens adopt healthy behaviors: no smoking, no drinking, attend prenatal clinic, eat properly (behavior change)
6. Teens have healthy babies (desired outcome)

Systems web showing possible influence linkages to a pregnant teenager

Program systems web showing possible institutional influences affecting pregnant teenagers:

Sustainable change: Systems dynamic reinforcing feedback loops
Using Different System Lenses to Understand a “particular” System

**Biologic System**
- Emergence
- Coordination/energy
- Structure, Process, Pattern
- Vitality

**Economic System**
- Inputs/Outputs
- Cost/Waste/Value/Benefits
- Customers/Suppliers

**Political System**
- Power
- Governance
- Citizenship
- Equity

**Sociologic System**
- Relationships
- Conversations
- Interdependence
- Loose-tight coupling
- Meaning/sense

**Anthropologic System**
- Values
- Culture/Milieu

**Mechanical / Physical System**
- Flow
- Temporal Sequencing
- Spatial Proximities
- Logistics
- Information

**Psychological System**
- Organizing
- Forces Field
- Psychological/Behaviour Setting

**SYSTEM DIMENSIONS**

Map Systems as Webs

Complex Dynamic Systems Configuration

HIV/AIDS Example
- Hits every system: health, family, social, religious, economic, political, community, international
- Requires multiple interventions on multiple fronts in all subsystems simultaneously
- Resulting reactions, interactions, consequences dynamic, unpredictable, emergent, and ever changing

Some system premises

**Systems neutrality:**
An observed system is functioning as observed for some reasons, fulfilling some functions.
In whose interests is a system functioning? Who benefits?

Healthy system
In a well-functioning system, no subsystem is operating at its maximum.
**Systems change**

During transitions from one system to another, things will get worse before they get better.

---

**Systems Dynamics**

Dynamic system interrelationships increase the likelihood of unintended consequences as systems change. Expect the unexpected.

---

**Challenges:**

Situation Recognition and Appropriate Evaluation Designs

---

The nature of **EXPERTISE:**

Situation Recognition

---

**Contingency-based Evaluation**

- Situational analysis & responsiveness
- Context sensitivity
- Clarify and focus on intended users: stakeholder analysis
- Clarify and focus on intended uses
- Methodological appropriateness
- Criteria for evaluating the evaluation: credibility, meaningfulness

---

**Conceptual Options**

- Simple
- Complicated
- Complex
21st Annual RTC Conference
Presented in Tampa, February 2008

**Types of Community Issues**

The Stacey Matrix

- **Certainty**
- **Agreement**

<table>
<thead>
<tr>
<th>Close to</th>
<th>Far from</th>
<th>Close to</th>
<th>Far from</th>
</tr>
</thead>
</table>

---

**Simple**

- Following a Recipe
  - The recipe is essential
  - Recipes are tested to assure replicability of later efforts
  - No particular expertise; knowing how to cook increases success
  - Recipes produce standard products
  - Certainty of same results every time

---

**Complicated**

- A Rocket to the Moon
  - Formulae are critical and necessary
  - Sending one rocket increases assurance that next will be ok
  - High level of expertise in many specialized fields + coordination
  - Rockets similar in critical ways
  - High degree of certainty of outcome

---

**Technically Complicated**

- Raising a Child
  - Technically Complicated
  - Experiment, coordinate expertise

---

**Socially Complicated**

- Build relationships, create common ground
Socially complicated situations pose the challenge of coordinating and integrating many players.

Environmental Initiatives
- Many different and competing stakeholders
- Diverse vested interests
- High stakes

Stakeholder Mapping

<table>
<thead>
<tr>
<th>High Interest/ Low Power</th>
<th>High Interest/ High Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>THE INVOLVED</td>
<td>THE PLAYERS</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>THE CROWD</td>
<td></td>
</tr>
<tr>
<td>Low interest/ Low Power</td>
<td></td>
</tr>
</tbody>
</table>

Know When Your Challenges Are In the Zone of Complexity

- System-Thinking
- Relationship-Building
- Collaboration
- Shared Knowledge
- Checking Around Drivers
- Minimum Specifications
- Multiple Actions
- Adaptability & Orgasm

Following a Recipe

- The recipe is essential
- Recipes are tested to assure replicability of later efforts
- No particular expertise: knowing how to cook increases success
- Recipes produce standard products
- Certainty of same results every time

Raising a Child

- Formulae have only a limited application
- Raising one child gives no assurance of success with the next
- Expertise can help but is not sufficient: relationships are key
- Every child is unique
- Uncertainty of outcome remains

Know When Your Challenges Are In the Zone of Complexity

- Zone of Complexity
- Agreement
- Certainty
- Chaos
- Massive Avoidance
**Simple**

Following a Recipe

- The recipe is essential
- Recipes are tested to ensure replicability of later efforts
- No particular expertise; knowing how to cook increases success
- Recipes note the quantity and nature of "parts" needed
- Recipes produce standard products
- Certainty of same results every time

---

**Complicate**

A Rocket to the Moon

- Formulas are critical and necessary
- Sending one rocket increases assurance that next will be ok
- High level of expertise in many specialized fields + coordination
- Separate into parts and then coordinate
- Rockets similar in critical ways
- High degree of certainty of outcome

---

**Complex**

Raising a Child

- Formulas have only a limited application
- Raising one child gives no assurance of success with the next
- Expertise can help but is not sufficient; relationships are key
- Can't separate parts from the whole
- Every child is unique
- Uncertainty of outcome remains

---

**The Frogtown Neighborhood Children’s Community Initiative in Saint Paul, Minnesota**

---

**Seeing Through A Complexity Lens**

"You don’t see something until you have the right metaphor to let you perceive it". Thomas Kuhn

---

**Complex Nonlinear Dynamics**

- Nonlinear: Small actions can have large reactions, “The Butterfly Wings Metaphor”
- Emergent: Self-organizing, Attractors
- Dynamic: Interactions within, between, and among subsystems and parts within systems can volatile, changing
- Getting to Maybe: Uncertainty, unpredictable, uncontrollable

---

**Major Sources of Uncertainty**

- Human irrationality: Behavioral Economics
- Different contexts
- Change in all its splendid manifestations

---

**New Direction # 7**

Methodological Flexibility & Creativity versus Methodological Rigidity
The Debate About Randomized Controls in Evaluation:
The Gold Standard Question

GOLD STANDARD:
METHODODOLOGICAL APPROPRIATENESS
not
Methodological orthodoxy or rigidity

Design Tension
Single Standard Hierarchy
vs
Situational Variation and Appropriateness

Creative Design Thinking
David P. Billington:*

“The goal of good design is to integrate efficiency, economy and elegance in a single design.”

* August 18, 2007, NY Times, A13

“One Bridge Doesn’t Fit All”

“As many have pointed out, the deadly bridge failure in Minneapolis was symptomatic of a system of bridges that will continue to corrode, crack and crumble if not maintained. But maintenance is not the only problem. We also need to design and build better bridges.”

Metaphor for evaluation design
“The Minneapolis collapse is hauntingly similar to the collapse in 1983 of another interstate highway bridge over the Mianus River in Connecticut. That disaster led to inspections of similar bridges, which found dangerous cracks from deferred maintenance....”

What the Mianus and Minneapolis bridges had in common was not just neglect. Both were the products of a design mentality in which engineers simply used a standard form, and often the same detailed features. Public bridges are all too often designed by anonymous teams, and the results can be seen on our highways.

Creative Challenge

Situational adaptability:
- Contingency-based evaluation
- Appropriateness
  -- Using standard forms of evaluation and
  -- Going beyond standard forms when appropriate and useful

Standard Evaluation Forms
1. Inadequate upfront utilization focus
2. Program/project as the unit of analysis
3. Linear logic models
4. Focus on findings use vs whole process
5. Individual outcomes focus vs systems change
6. Preference for quantitative data & RCTs as the methodological Gold Standard
7. Static designs

Frances Westley, Brenda Zimmerman, Michael Q. Patton
Random House Canada,

Complex Situations
- Highly emergent (difficult to plan and predict)
- Highly dynamic, rapidly changing
- Relationships are non-linear & interdependent rather than simple (linear cause-effect)
Complex Interdependencies

Insert action into the system

Contingency-based Developmental Evaluation

Improvement versus Development

Beyond just Summative and Formative
Beyond Static Accountability Models

Example of an emergent option:

Developmental Evaluation

DEVELOPMENTAL EVALUATION DEFINED
Evaluation processes, including asking evaluative questions and applying evaluation logic, to support program, product, staff and/or organizational development. The evaluator is part of a team whose members collaborate to conceptualize, design and test new approaches in a long-term, on-going process of continuous improvement, adaptation and intentional change. The evaluator’s primary function in the team is to elucidate team discussions with evaluative questions, data and logic, and facilitate data-based decision-making in the developmental process.

CONTRASTS

Traditional evaluations... Testing models
• Render definitive judgments of success or failure

Complexity-based, Developmental Evaluation...
• Supporting innovation and adaptation

Traditional Evaluation... Developmental Evaluation...
• Provide feedback, generate learnings, support direction or affirm changes in direction in real time

• Measure success against predetermined goals

• Develop new measures and monitoring mechanisms as goals emerge & evolve
21st Annual RTC Conference
Presented in Tampa, February 2008

**Traditional Evaluation...**
- Evaluator external, independent, objective

**Developmental Evaluation...**
- Evaluator part of a team, a facilitator and learning coach bringing evaluative thinking to the table, supportive of the organization’s goals

**Traditional Evaluation...**
- Evaluator determines the design based on the evaluator’s perspective about what is important. The evaluator controls the evaluation.

**Developmental Evaluation...**
- Evaluator collaborates with those engaged in the change effort to design an evaluation process that matches philosophically and organizationally.

**Traditional Evaluation...**
- Design the evaluation based on linear cause-effect logic models

**Developmental Evaluation...**
- Design the evaluation to capture system dynamics, interdependencies, and emergent interconnections

**Traditional Evaluation...**
- Aim to produce generalizable findings across time & space.

**Developmental Evaluation...**
- Aim to produce context-specific understandings that inform ongoing innovation

**Traditional Evaluation...**
- Accountability focused on and directed to external authorities and funders.

**Developmental Evaluation...**
- Accountability centered on the innovators’ deep sense of fundamental values and commitments – and learning.

**Traditional Evaluation...**
- Learning to respond to lack of control and stay in touch with what’s unfolding
- And thereby respond strategically.
Traditional Evaluation...
• Evaluation often a compliance function delegated down in the organization

Developmental Evaluation...
• Evaluation a leadership function:
  Reality-testing, results-focused, learning-oriented leadership

Traditional Evaluation...
• Evaluation supports hunger for learning.

Developmental Evaluation...
• Evaluation engenders fear of failure.

Conditions
• High innovation
• Development
• High uncertainty
• Dynamic
• Emergent
• Systems change

SenseMaker software
• Dave Snowden, Founder of Cognitive Edge, former Director of Knowledge Management at IBM
• SenseMaker can code and map 95,000 stories in 24 hours
• See the world as others see it; anti-terror applications.
• See the quantitative patterns in the metadata with qualitative context and meaning

New Direction # 8
Infusing evaluative thinking as a primary type of evaluation process use.

Capacity-building as an evaluation focus.

Thinking about:
The role can evaluation play with complex dynamic innovations....
And the beat goes on...

Evaluation as an ever-evolving field

References

• *Getting to Maybe: How the World Is Changed?* Frances Westley, Brenda Zimmerman, Michael Q. Patton