Major Directions, Issues, and Trends in Evaluation

with

Michael Quinn Patton

June 16, 2009
10. Evaluation as an International and Cross-Cultural Transdiscipline and Profession

• National/regional associations and societies (70+)
• IOCE and IDEAs
• Toronto, 3rd International Evaluation Conference, 2005: General Romeo Dallaire's "Speaking truth to Power" Award from AEA & CES
• Evaluators' cultural competence and evaluation relevance as issues
• International adoption of standards
Evaluation Standards

- **Utility** – ensure relevance & use
- **Feasibility** – realistic, prudent, diplomatic & frugal
- **Propriety** – ethical, legal, respectful
- **Accuracy** – technically adequate to determine merit or worth

For the full list of Standards:

[www.wmich.edu/evalctr/checklists/standardschecklist.htm](http://www.wmich.edu/evalctr/checklists/standardschecklist.htm)
9. Utilization focus:
Intended Use by Intended Users

From...

Audiences to...

Stakeholders to...

Primary Intended Users

Connotative differences?
Evaluators’ Views on Practice

• As part of a review of developments over the first ten years of the American Evaluation Association (AEA), Preskill and Caracelli (1997) conducted a survey of members of AEA's Topical Interest Group on Use. They found that 85% rated as extremely or greatly important "identifying and prioritizing intended users of the evaluation" (p. 216). The only item eliciting higher agreement (90%) was the importance of "planning for use at the beginning of the evaluation." Preskill and Caracelli also found that 80% of survey respondents agreed that evaluators should take responsibility for involving stakeholders in the evaluation processes.
Evaluators’ Views Cont.

• Fleischer (2007) asked the same question on a replication survey of American Evaluation Association members in 2006 and found that 98% supported the importance of focusing on intended use by intended users. In rating the importance of eight different evaluation approaches, "user-focused" evaluation was rated highest.
Recently, Cousins and Shulha (2006) reviewed a great volume of research on utilization of evaluation and knowledge found that "both social scientists and evaluators are learning that attention to the characteristics of knowledge users is a potent way to stimulate the utilization of findings" (p. 273).
Taking user *cognitive style* into account

- Lawrence Lynn of Harvard argues in his authoritative and still-relevant case book on policy analysis that a major craft skill needed by policy and evaluation analysts is the ability to understand and make accommodations for specific decision maker's cognitive style and other personal characteristics.
Style difference examples

• Jerry Brown, former Governor of California, liked policy analyses framed as a debate—thesis, antithesis—because he had been trained in the Jesuitical style of argument.

• President Ronald Reagan liked *Reader’s Digest* style stories and anecdotes.

• President Jimmy Carter wanted details (engineering mind)

• Former Secretary of Defense and World Bank President Robert MacNamara was a systems analysis and demanded at least five alternatives for any problem.
8. Increased political attention to accountability:

Evaluation language as political rhetoric

• Targets, benchmarks, and milestones, e.g., MDGs, Kyoto Agreement
• Best practices, Promising practices
• Results-based management, Results-based decision-making, Results-based mandates
• Accountability
• GPRA, PART, NCLB
• Chief Performance Officer, White House
Challenges of Misuse and Abuse: Speaking Truth to Power

Eleanor Chelimsky:

"Telling the truth to people who may not want to hear it is, after all, the chief purpose of evaluation"

(Chelimsky, 1995: 54)
7. Long-time tensions endure

• *Competing purposes*: accountability, learning, management, improvement (formative), knowledge generation, summative judgment, validating models

• *Relationship and role issues*: independence, mutual understanding, trust, respect

• *Credible data*: Quantitative vs qualitative data, mixed methods

• *Priority primary intended users*: Policy makers, administrators, program staff, beneficiaries, interest groups, advocates, general public
6. **Evaluation capacity-building:**
Attention to the capacities and essential skills needed to conduct and use evaluations

- Evaluation and Organizational Development connecting & cross-fertilizing
- Changes in the evaluation unit of analysis:
  - Strategies, Policies, Initiatives, Organizations, Collaborations
5. *Process use*: Beyond findings use

**Key distinction:**

*Process Evaluation* versus *Process Use*
Process Use Defined

Process use refers to and is indicated by individual changes in thinking and behavior, and program or organizational changes in procedures and culture, that occur among those involved in evaluation as a result of the learning that occurs during the evaluation process. Evidence of process use is represented by the following kind of statement after an evaluation: "The impact on our program came not just from the findings, but from going through the thinking process that the evaluation required."
Examples of Process Uses

• Enhancing shared understandings, e.g., logic model development
• Focusing programs: What gets measured gets done
• Effects on organizational culture: accountability culture, results-orientation
• Capacity-building for those involved, deepening evaluative thinking
• Program and organizational development, e.g., evaluability assessments
Culture of Accountability Effects

“Western aid workers seem to spend most of their time drawing up flow charts for each other. They’re so worried about their inspectors general that they can’t really immerse themselves in the messy world of local reality.”

• David Brooks, *NY Times*, March 27, 2009

http://www.nytimes.com/2009/03/27/opinion/27brooks.html?_r=1
4. From Studies to Streams: Real time evaluation

• Rapid feedback
• Timeliness
• Ongoing assessments
• Continuous information flows, updates
• Use of the Internet and new technologies

3. Challenges in establishing attribution

- Gold Standard debate and confusion
The Attribution Debate:

• The Gold Standard Question
GOLD STANDARD:

METHODOLOGICAL APPROPRIATENESS

not

Methodological orthodoxy or rigidity
Attribution approaches

- Mixed methods: triangulation
- Methodological appropriateness
- *General elimination* method
- Preponderance of evidence criterion

*See appendix for more details*
2. More sophisticated recommendations, reports and graphics

Reports: Ever greater attention to Executive Summaries and Briefings

Importance of focusing on intended use by intended users
The Extreme Presentation™ Method

Ten Steps to Presentation Impact

1. Audience
2. Objectives
3. Problem-Solution
4. Evidence
5. Anecdotes
6. Sequencing
7. Charts
8. Layouts
9. Stakeholders
10. Measurement

Politics and Metrics

Impact

© 2008 A. Abela
Powerful Graphics: Visual Thinking

• Mike Hendricks regularly trains evaluators on reporting and he asserts emphatically: “Evaluators have got to learn graphics. I’m amazed at how bad the charts and graphics are that I see in reports. You can’t emphasize it too much. Reporting means GRAPHICS! GRAPHICS! GRAPHICS!” This involves "visible thinking" which includes causal mapping and other data displays (Bryson, Ackermann, Eden and Finn 2004).
Significance

Fred Mosteller, the great applied statistician, was fond of saying that he did not care much for statistically significant differences. He was more interested in *interocular differences* - the differences that hit us between the eyes.

Otherwise, we risk transauditory impacts....
Historical example:

Through her work as a nurse in the Crimean War, Florence Nightingale was a pioneer in establishing the importance of sanitation in hospitals. She meticulously gathered data on relating death tolls in hospitals to cleanliness, and, because of her novel methods of communicating this data, she was also a pioneer in applied statistics.
Table showing the Estimated Average Monthly Strength of the Army; and the deaths and Annual Rate of Mortality per 1000 in each month, from April 1854 to March 1856 (inclusive), in the Hospitals of the Army in the East.

<table>
<thead>
<tr>
<th>Month</th>
<th>Average size of army</th>
<th>Deaths Zymotic Diseases</th>
<th>Wounds &amp; Injuries</th>
<th>All other causes</th>
<th>Annual rate of mortality per 1000</th>
<th>Zymotic Diseases</th>
<th>Wounds &amp; Injuries</th>
<th>All other causes</th>
<th>All other causes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr 1854</td>
<td>8571</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>6.2</td>
<td>0</td>
<td>4.6</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>May 1854</td>
<td>23333</td>
<td>12</td>
<td>0</td>
<td>9</td>
<td>6.2</td>
<td>0</td>
<td>4.6</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Jun 1854</td>
<td>28333</td>
<td>11</td>
<td>0</td>
<td>6</td>
<td>4.7</td>
<td>0</td>
<td>2.5</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Jul 1854</td>
<td>28722</td>
<td>359</td>
<td>0</td>
<td>23</td>
<td>150.0</td>
<td>0</td>
<td>9.6</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Aug 1854</td>
<td>30246</td>
<td>828</td>
<td>1</td>
<td>30</td>
<td>328.5</td>
<td>0.4</td>
<td>11.9</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Sep 1854</td>
<td>30290</td>
<td>788</td>
<td>81</td>
<td>70</td>
<td>312.2</td>
<td>32.1</td>
<td>27.7</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Oct 1854</td>
<td>30643</td>
<td>503</td>
<td>132</td>
<td>128</td>
<td>197.0</td>
<td>51.7</td>
<td>50.1</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Nov 1854</td>
<td>29736</td>
<td>844</td>
<td>287</td>
<td>106</td>
<td>340.6</td>
<td>115.8</td>
<td>42.8</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Dec 1854</td>
<td>32779</td>
<td>1725</td>
<td>114</td>
<td>131</td>
<td>631.5</td>
<td>41.7</td>
<td>48.0</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Jan 1855</td>
<td>32393</td>
<td>2761</td>
<td>83</td>
<td>314</td>
<td>1022.8</td>
<td>30.7</td>
<td>120.0</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Feb 1855</td>
<td>30919</td>
<td>2120</td>
<td>42</td>
<td>361</td>
<td>822.8</td>
<td>16.3</td>
<td>140.1</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Mar 1855</td>
<td>30107</td>
<td>1205</td>
<td>32</td>
<td>172</td>
<td>480.3</td>
<td>12.8</td>
<td>68.6</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Apr 1855</td>
<td>32252</td>
<td>477</td>
<td>48</td>
<td>57</td>
<td>177.5</td>
<td>17.9</td>
<td>21.2</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>May 1855</td>
<td>35473</td>
<td>508</td>
<td>49</td>
<td>37</td>
<td>171.8</td>
<td>16.6</td>
<td>12.5</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Jun 1855</td>
<td>38863</td>
<td>802</td>
<td>209</td>
<td>31</td>
<td>247.6</td>
<td>64.5</td>
<td>9.6</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Jul 1855</td>
<td>42647</td>
<td>382</td>
<td>134</td>
<td>33</td>
<td>107.5</td>
<td>37.7</td>
<td>9.3</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Aug 1855</td>
<td>44614</td>
<td>483</td>
<td>164</td>
<td>25</td>
<td>129.9</td>
<td>44.1</td>
<td>6.7</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Sep 1855</td>
<td>47751</td>
<td>189</td>
<td>276</td>
<td>20</td>
<td>47.5</td>
<td>69.4</td>
<td>5.0</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Oct 1855</td>
<td>46852</td>
<td>128</td>
<td>53</td>
<td>18</td>
<td>32.8</td>
<td>13.6</td>
<td>4.6</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Nov 1855</td>
<td>37853</td>
<td>178</td>
<td>33</td>
<td>32</td>
<td>56.4</td>
<td>10.5</td>
<td>10.1</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Dec 1855</td>
<td>43217</td>
<td>91</td>
<td>18</td>
<td>28</td>
<td>25.3</td>
<td>5.0</td>
<td>7.8</td>
<td>13</td>
<td>13</td>
</tr>
</tbody>
</table>
The Areas of the blue, red, & black wedges are each measured from the centre as the common vertex.
The blue wedges measured from the centre of the circle represent area for area the deaths from Preventable or Mitigable Zymotic diseases, the red wedges measured from the centre the deaths from wounds, & the black wedges measured from the centre the deaths from all other causes.
The black line across the red triangle in Nov’ 1854 marks the boundary of the deaths from all other causes during the month.
In October 1854, & April 1855, the black area coincides with the red, in January & February 1856 the blue coincides with the black.
The entire areas may be compared by following the blue, the red & the black enclosing lines.

☑️ Preventable or mitigable zymotic diseases
☑️ Wounds
☑️ Other causes
Chart Suggestions—A Thought-Starter

Comparison

What would you like to show?

Relationship

Distribution

Composition

32
Fine-tuning Recommendations

• Controversy
• Data-based
• Futuring perspective
• Understanding the situation:
  Situationally appropriate recommendations
The nature of **EXPERTISE**
Conceptual Options

- Simple
- Complicated
- Complex
Types of Community Issues
The Certainty/Agreement Matrix

![Diagram of the Certainty/Agreement Matrix]
Simple situations

Agreement

Certainty

Simple

Plan, control

Close to

Far from

Close to

Far from
Simple (Known arena of action)

- Tight, centralized connections.
- Can identify and make sense of patterns.
- Linear cause and effect.
- Best practices identifiable within the current context (which of course may not be self-evident or known to others – hence importance of context).
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
Recommendations in simple space: substantive, action-focused recommendations
Technically Complicated
Complicated (Knowable arena)

• Relationships are looser but still clustered around a central core.
• Cause and effect is dynamic, multi-dimensional, and enmeshed in system relationships.
• System relationships can be modelled and understood.
• Expertise and coordination needed.
**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

**Complex**
Raising a Child
Recommendations in technically complicated space

- Improving system coordination
- Understanding system elements and their functions
- Ways of dealing with complications
Socially Complicated

Agreement

Close to

Far from

Certainty

Close to

Far from

Socially Complicated

Build relationships, create common ground

Simple

Plan, control

Technically Complicated

Experiment, coordinate expertise
Socially complicated

Implementing human rights agreements, like gender equity or outlawing child labor

Environmental Initiatives

- Many different and competing stakeholders
- Diverse vested interests
- High stakes
Socially complicated situations pose the challenge of coordinating and integrating many players.
Recommendations in socially complicated space

- Improving system coordination
- Understanding key participants and their roles
- Ways of dealing with relationships
- Improving collaboration
Know When Your Challenges Are In the Zone of Complexity

- **Socially Complicated**: Build relationships, create common ground
- **Simple**: Plan, control
- **Technically Complicated**: Experiment, coordinate expertise

**Agreement**

**Certainty**

- **Close to**
- **Far from**

**Systems Thinking**
**Relationship Building**
**Collaboration**
**Good Enough Vision**
**Chunking Around Drivers**
**Minimum Specifications**
**Multiple Actions**
**Adaptability & Organic**
Complex

• Centre is loosely connected to network.
• Cause effect difficult to track; nonlinear, interdependent relationships
• Highly context dependent.
• Outcomes emergent, not predictable.
### 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

- 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
- 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

### Complex
Raising a Child

- 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
- 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

- Close to Agreement, Close to Certainty: Simple
  - Plan, control
- Far from Agreement, Close to Certainty: Technically Complicated
  - Experiment, coordinate expertise
- Close to Agreement, Far from Certainty: Socially Complicated
  - Build relationships, create common ground
- Far from Agreement, Far from Certainty: Chaos
  - Massive Avoidance

Zone of Complexity
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
Recommendations under conditions of complexity

- Key factors to monitor
- What to pay attention to
- Alert to unanticipated consequences, side effects, and emergence
- Rapid responses
- Flexibility
Classic mistake

• Recommending treating the complex as simple

Wise executives tailor their approach to fit the complexity of the circumstances they face.
<table>
<thead>
<tr>
<th>COMPLEX</th>
<th>EMERGENT</th>
<th>KNOWABLE</th>
<th>COMPLICATED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cause and effect are only coherent in retrospect and do not repeat</td>
<td>Cause and effect separated over time and space</td>
<td>Analytical/Reductionist</td>
<td></td>
</tr>
<tr>
<td>Pattern management</td>
<td>Scenario planning</td>
<td>Systems thinking</td>
<td></td>
</tr>
<tr>
<td>Perspective filters</td>
<td>Systems thinking</td>
<td>Sense-Analyze-Respond</td>
<td></td>
</tr>
<tr>
<td>Complex adaptive systems</td>
<td>Complex adaptive systems</td>
<td>Complex adaptive systems</td>
<td></td>
</tr>
<tr>
<td>Probe-Sense-Respond</td>
<td>Probe-Sense-Respond</td>
<td>Probe-Sense-Respond</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TURBULENT</th>
<th>KNOWN</th>
<th>SIMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cause and effect relationships perceivable</td>
<td>Cause and effect relations repeatable, perceivable and predictable</td>
<td>Stability-focused intervention</td>
</tr>
<tr>
<td>Stability-focused intervention</td>
<td>Legitimate best practice</td>
<td>Enactment tools</td>
</tr>
<tr>
<td>Enactment tools</td>
<td>Standard operating procedures</td>
<td>Crisis management</td>
</tr>
<tr>
<td>Crisis management</td>
<td>Process reengineering</td>
<td>Act-Sense-Respond</td>
</tr>
<tr>
<td>Act-Sense-Respond</td>
<td>Sense-Categorize-Respond</td>
<td>Act-Sense-Respond</td>
</tr>
</tbody>
</table>
1. Systems Thinking and Complexity Frameworks

• Parts are interdependent such 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
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

- Prenatal program staff
- Her parents & other family members
- Teachers/other adults
- Child's father & peers
- Her peer group

Young pregnant woman's attitudes & behaviors
Program systems web showing possible institutional influences affecting pregnant teenagers:

**Other Systems**
- welfare
- legal
- nutrition programs
- transportation
- child protection
- media messages

**Context factors**
- politics
- economic incentives
- social norms
- culture
- music

- Prenatal program
- Prenatal Clinic and Hospital Outreach
- Youth Culture
- Other community-based youth programs
Using Different System Lenses to Understand a “particular” System

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

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

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

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

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

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

**Information System**
- Access
- Speed
- Fidelity/utility
- Privacy/security
- Storage

**Psychological System**
- Organizing
- Forces Field
- Ecological/Behaviour Settings

**SYSTEM DIMENSIONS**
Map Systems as Webs

Source: Digital Capital: Harnessing the Power of Business Webs, By Don Tapscott, David Ticoll and Alex Lowy
Evaluation as an evolving field

From the beginning...

to the present...

and into the future...
Table discussion questions

1. What challenges do you face internally in your I&E unit?
   a. How do you deal with those challenges?
2. What challenges do you face interacting with key external stakeholders?
   a. How do you deal with those challenges?
3. What are your lessons learned about recommendations that you have made?
Resources:
(1) Chart alternatives graphic
www.ExtremePresentation.com
© 2009 A. Abela

(2) Florence Nightingale examples:
http://understandinguncertainty.org/node/214

Reference: