Designing useful evaluations that actually get used

with

Michael Quinn Patton
Research-based Approach

• Follow-up study of the use of 20 federal health evaluations
• Canadian research on evaluation use
• U-FE now widely used among professional evaluators

(see appendix slides)
Factors Affecting Use

➢ Table Exercise
Evaluation and Research: Same or different? And why?
Foundational Premise:

✓ Research and evaluation are different – and therefore…

✓ Evaluated by different standards.
Evaluation Standards

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

For the full list of Standards:

www.wmich.edu/evalctr/checklists/standardschecklist.htm
Evaluation Models

U-FE is one among many....
Cultural context

• Emergence and meaning of the jester
Behold the "stake-holder"!
I HIT MY GOAL DEAD CENTER
I hit my goal... dead center!
Changing societal and political context for evaluation
Evaluation Hits Popular Culture

“When I grow up, I want to be”...

“...a results-oriented cowboy.”

“...an evidence-based fire fighter.”

“...a performance-monitoring, accountability-driven movie star.”
Lessons

What we’ve learned about designing useful evaluations that actually get used.

A case example…
Example

Caribbean Agricultural Extension Project

U.S. AID
European Development Bank
Caribbean Development Bank
University of the West Indies
Big Ten Universities (MUCIA)
CARDI
Overview

• 10 year project in 3 phases
• Phase 1, January 1980 to December 1982
• Mid-term and end-of-phase evaluations planned and budgeted
• 10 Caribbean countries
Utilization Question

When will evaluation results be needed to contribute to the summative decision about whether to fund Phase 2?

When do you think?
When will evaluation results be needed to contribute to the summative decision about whether to fund Phase 2?

State Department budget to Congress

✓ Caribbean Budget to State Department
✓ Barbados Regional Office to Wash DC and Caribbean Region Office
✓ Agriculture program to Caribbean
Evaluation Questions

• Are outstanding extension agents making a significant contribution to improved farming for small farmers?
  
  Answer “no”: End project

  Answer “yes”: Consider next question…

• If so, is it worth training more such agricultural extension agents?
An ancient example

• Group exercise:

What lessons about making evaluation useful do you extract from this example?
Lessons?
Utilization-Focused Evaluation (U-FE)

A decision-making framework for enhancing the utility and actual use of evaluations.
U-FE begins with the premise that evaluations should be judged by their utility and actual use. Therefore, evaluators should facilitate the evaluation process and design any evaluation with careful consideration of how everything that will be done, from beginning to end, will affect use.
USE

• Take use seriously by evaluating use, the source of our own accountability and ongoing learning/professional development
• Different from dissemination
• Different from producing reports
• Groundwork laid and expectations set at the beginning
• Doesn’t happen naturally or automatically
Goal of U-FE

Intended Use by Intended Users
Intended Evaluation Users

From…
Audiences to…
Stakeholders to…

Primary Intended

Connotative differences?
We're sure this evaluation will completely meet ALL the information needs of absolutely EVERYONE.

Would you buy a used car from this evaluator?
Stakeholder Mapping

**THE INVOLVED**

- High Interest/
  - Low Power

**THE PLAYERS**

- HIGH Interest/
  - High Power

**THE CROWD**

- Low Interest/
  - Low Power

**CONTEXT SETTERS**

- Low Interest/
  - High Power
High interest stakeholders

Low power stakeholders
Support and enhance their capacity to be involved, especially when they may be affected by findings, as in the case of program participants. Their involvement increases the diversity of the evaluation.

High power stakeholders
High potential as primary intended users. These are often key "players" who are in a prime position to affect use, including using it themselves as well as drawing the attention of others.

Low interest stakeholders
Inform them about the evaluation and its findings. Controversy can quickly turn this amorphous "crowd" of general public stakeholders into a very interested mob.

Need to cultivate their interest and be alert in case they pose barriers to use through their disinterest. They are "context setters."

Michael Quinn Patton
June 14, 2010
Research Base

Studies of evaluation use have identified primary user identification and involvement as a critical factor contributing to evaluation use.
Goal of U-FE:

Intended Use by Intended Users
Different Evaluation Purposes

• For making judgments
  Commonly called *summative evaluations*:

• For improving programs
  Commonly called *formative evaluations*

• For ongoing development
  Sometimes called *developmental evaluations*

• For knowledge building
  Meta-evaluation, lessons learned, effective practices
Additional uses

• Accountability

• Monitoring (M & E) integrated
Tensions

Different intended uses serve different purposes and, typically, different intended users.

Thus the need to FOCUS and manage tensions between and among different purposes.
Premises of Utilization Focused Evaluation

• No evaluation should go forward unless and until there are primary intended users who will use the information that can be produced
• Primary intended users are involved in the process
• Evaluation is part of initial program design - The primary intended users want information to help answer a question or questions.
• Evaluator’s role is to help intended users clarify their purpose and objectives.
• Make implications for use part of every decision throughout the evaluation – the driving force of the evaluation process.
Use Distinctions

• **Instrumental Use**
• **Conceptual Use**
• **Misuse**
Campbell's Law

"The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor."

Campbell (1975)
Beyond Misevaluation

Essential evaluator competencies

• Beyond technical competence
• Professional knowledge (e.g., standards & guiding principles)
• Interpersonal, communications, facilitation, & political skills
• Cultural competence
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)
ROMÉO DALLAIRE
FORCE COMMANDER OF THE UN ASSISTANCE MISSION TO RWANDA, 1993-1994
SHAKE HANDS WITH THE DEVIL
THE FAILURE OF HUMANITY IN RWANDA
“EXTRAORDINARY, WRENCHING LYRIC POWER.”—MONTREAL GAZETTE
WINNER
2004
GOVERNOR GENERAL’S
LITERARY AWARD FOR
NONFICTION
Canada Council
for the Arts
FOREWORD BY SAMANTHA POWER
SPEAKING TRUTH TO POWER

Michael Quinn Patton
June 14, 2010
Evaluation Use Factors

• Evaluation & politics
• Evaluation & leadership
• Evaluation & organizational culture
• Evaluation & systems change
• Evaluation & the social science of change
New Direction

Systems thinking & complexity science
as frameworks for conceptualizing interventions
(replacing linear logic models)

Developmental Evaluation

Wise executives tailor their approach to fit the complexity of the circumstances they face.
COMPLEX
Cause and effect are only coherent in retrospect and do not repeat
Pattern management
Perspective filters
Complex adaptive systems
Probe-Sense-Respond

KNOWABLE
Cause and effect separated over time and space
Analytical/Reductionist
Scenario planning
Systems thinking
Sense-Analyze-Respond

CHAOS
No cause and effect relationships perceivable
Stability-focused intervention
Enactment tools
Crisis management
Act-Sense-Respond

KNOWN
Cause and effect relations repeatable, perceivable and predictable
Legitimate best practice
Standard operating procedures
Process reengineering
Sense-Categorize-Respond
Wise evaluators tailor their approach to fit the complexity of the circumstances they face
The Challenge:

Matching the evaluation design to the evaluation’s purpose, resources, and timeline to optimize use.

Focus: Intended use by intended users
Leaders as Primary Intended Users

Reality-Testing, Results-Oriented, Learning-Focused Leadership
Four Leadership Functions

1. Create and support a results-oriented, reality-testing, learning-focused culture.
Reality-testing
Results-oriented
BREVITY BY GUY & RODD

I THINK I CAN, I THINK I CAN...
The importance of Organizational Culture in determining use
Situation Analysis

What are the critical organizational culture factors that affect evaluation use?

- Enablers/facilitating factors
- Barriers to use
Leadership Functions

1. Create and nurture a results-oriented, reality-testing culture.

2. Lead in deciding what outcomes to commit to and hold yourselves accountable for.
Outcomes are an expression of Values: E-valuation
1. Create and nurture a results-oriented, reality-testing culture.

2. Lead in deciding what outcomes to commit to and hold yourselves accountable for.

3. Make measurement of outcomes thoughtful, meaningful and credible.
Use Distinctions

Findings Use

and

Process Use
Process Use

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 so much from the findings but from going through the thinking process that the evaluation required."
Process Uses

• Enhancing shared understandings
• Focusing programs: What gets measured gets done
• Supporting and reinforcing the program intervention, e.g., feedback for learning
• Capacity-building for those involved, deepening evaluative thinking
• Program and organizational development, e.g., evaluability assessments
Do you have anything to declare?

Yes. Evaluation is a way of thinking.

They're only looking for findings.

Michael Quinn Patton
June 14, 2010
New Direction

Infusing *evaluative thinking* as a primary type of process use.

IDRC example

Capacity-building as an evaluation focus of process use.
The importance of Organizational Culture in supporting use & learning
• Evaluation capacity-building as a priority to support use
Basic evaluation literacy

• Know the evaluation standards
• Know how to apply the standards in the actual conduct of evaluations
• Understand different potential uses and their implications methodologically and procedurally
• Understand how to identify and work with primary intended users
• Have evaluators with essential skills
1. Create and nurture a results-oriented, reality-testing culture.

2. Lead in deciding what outcomes to commit to and hold yourselves accountable for.

3. Make measurement of outcomes thoughtful, meaningful and credible.

4. Use the results -- and model for others serious use of results.
Use Scenarios

• Decisions and actions
• Priority uses
• Realistic timelines to increase utility and actual use
• Interpretation engagement
• Recommendation options
Walking the Talk:
Reinforcing a Reality-testing, Results-oriented, Learning-committed organizational culture
Developmental Evaluation:
Applying Complexity Concepts to Enhance Innovation and Use

Michael Quinn Patton
I evaluate; therefore, I am.
And the beat goes on…

Evaluation as an ever-evolving field
Reference

Utilization-Focused Evaluation, 4th ed, Michael Quinn Patton

http://www.sagepub.com/booksProdDesc.nav?
prodId=Book229324
Appendix

These are slides as resources in answering questions and going deeper into some issues. Time may not permit covering these during the Institute workshop.
Research on U-FE use among evaluators:

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*
Fleischer (2007) asked the same question on a replication survey of American Evaluation Association members in 2006 and found that 98% agreed with this assertion.

In rating the importance of eight different evaluation approaches, "user-focused" evaluation was rated highest.

Stakeholder involvement in evaluations has become accepted practice among evaluation professionals.
Summary Lessons on Useful Evaluation

1. Clearly identify primary intended users
2. Clearly identify primary intended uses
   Goal: Intended use by intended users
3. Negotiate FOCUS -- get agreement on criteria
4. Establish a clear ACTION framework
5. Distinguish empirical questions from value questions
6. Select methods appropriate to the question
7. Facilitate actual use of the findings
Classic S-shaped curve
Rogers’ adopter categories

- **Innovators**: Innovators are the first individuals to adopt an innovation. Innovators are willing to take risks, and have closest contact to scientific sources and interaction with other innovators.

- **Early Adopters**: This is second fastest category of individuals who adopt an innovation. These individuals have the highest degree of opinion leadership among the other adopter categories. Early adopters are typically younger in age, have a higher social status, advanced education, and are more socially forward than late adopters.
Adopter categories

• **Early Majority**: Individuals in this category adopt an innovation after a varying degree of time. This time of adoption is significantly longer than the innovators and early adopters. Early Majority tend to be slower in the adoption process, have above average social status, contact with early adopters, and show some opinion leadership.
Adopter categories

• **Late Majority**: Individuals in this category will adopt an innovation after the average member of the society. Late Majority are typically skeptical about an innovation, have below average social status, in contact with others in late majority and very little opinion leadership.

• **Laggards**: Individuals in this category are the last to adopt an innovation. Unlike some of the previous categories, individuals in this category show little to no opinion leadership. These individuals typically have an aversion to change-agents and tend to be advanced in age. Laggards typically tend to be focused on “traditions”, and very little to no opinion leadership.
HYMN

to

EVALUATION

USE

Sung to the tune of

Auld Lang Syne
May all e-valu-a-tions done
Be useful as they should
They tell us how to separate
What is poor from what is good.

We gather data near and far
To see what we can learn
The findings help us to decide
What to keep and what to burn.
There comes a time for each of us
When doubts may give us pause
We wonder what results will show
Will the world see naught but flaws

But be assured there's naught to fear
If learn-ing is what you seek
Let outcomes guide your every move
Listen to the data speak.
TRIANGUALTED LEARNING FRAMEWORK

ACTION

BELIEFS/VALUES  KNOWLEDGE

Michael Quinn Patton
June 14, 2010
Premises

• Action flows from some combination of beliefs/values and knowledge.
• Informed action has a strong knowledge base.
• By testing and evaluating beliefs, we build knowledge, thereby making our actions more informed and empirically based.
• People have varying predilections that lean more toward one of these dimensions than others, e.g., more action-oriented, more values-oriented, or more knowledge-oriented.
• All three styles are valuable and needed for sustainable program, organizational and/or community development.
• In short, **ACTIONS** flow from some combination of **IDEAS** (THEORY), **BELIEFS** (MISSION/VISION/VALUES) and **KNOWLEDGE** (EVIDENCE).

• A **Learning Organization**, over time, moves more action to a strong knowledge base.
Alignment and consistency support meaningful, high-impact and Sustainable results.

Images of Alignment Between Actions, Beliefs and Knowledge

Non-aligned:
Inconsistencies between actions, beliefs and knowledge

Aligned
Consistency between actions, beliefs and knowledge
Utilization-Focused

Methods Decisions
The Debate About Randomized Controls in Evaluation:

• The Gold Standard Question
New Direction

The morphing of the paradigms debate (qualitative vs quantitative) into the Gold Standard debate (randomized control trials as the alleged "gold standard" for impact evaluation)
• No single design or method is universally “strongest”
• Multiple ways of establishing causality
• Dangerous to privilege one method, creates perverse incentives
GOLD STANDARD:

METHODOLOGICAL APPROPRIATENESS

not

Methodological orthodoxy or rigidity