

**Options for Evaluating
Youth Voluntary Service:
Diverse Methods and
Standards of Evidence.**
by

Michael Quinn Patton

Utilization-Focused Evaluation

May 8, 2008

mqpatton@prodigy.net

Context:

Youth service was cited as a “Promising but unproven” strategy for youth development in the 2007 WDR.

- What does it mean to “prove” youth service as a strategy for youth development?
- Can this be proven?

Philosophy of Science

*To Karl Popper, falsifiability,
not verifiability,
is the essence of science.*

You can't prove.

You can only disprove.

“Proof” has become a political construct

- Rhetoric of evidence-based practice
- Best practices
- Proven practices

Debate about evidence for
“PROOF”:

Experimental designs (RCTs)

vs.

Preponderance of evidence
standard

Position of Professional Evaluation Associations

- AEA position (Nov 4, 2003):
Randomized control group trials (RCTs) not the only studies capable of generating understandings of causality.
- Many methods capable of demonstrating scientific rigor. Methods.”
<http://www.eval.org/doestatement.htm>

European Evaluation Society position

As context, the EES statement is in response to strong pressure from some interests advocating for “scientific” and “rigorous” impact of development aid, where this is defined as primarily involving RCTs. This debate has the potential to influence the future direction of evaluation—not only with respect to development but potentially in other areas as well. EES however deplores one perspective currently being strongly advocated: that the best or only rigorous and scientific way of doing so is through randomized controlled trials (RCTs). In contrast, the EES supports multi-method approaches to IE [impact evaluation] and does not consider any single method such as RCTs as first choice or as the “gold standard.”

www.europeanevaluation.org/news?newsId=1969406

In 2007, the Network of Networks on Impact Evaluation – **NONIE** -- was established by international evaluation offices representing more than 100 United Nations, World Bank, and other development organizations, plus representatives from developing countries and various regional and global organizations. That group drafted a document providing guidance for conducting impact evaluations in developing countries.

NONIE's draft statement emphasizes the importance of methodological diversity and appropriateness in support of rigor, and warned against designating any single design as a gold standard.

Children & Youth, The World Bank

Volume II, Number 5, June 2007

Elements of effective impact evaluation

- Isolate the effect of a youth development program on a given outcome, or to test the optimal combination of interventions in different contexts.
- Impact evaluation helps us understand “what is the effect of X on Y?” For example: what is the effect of a youth training program on employment? Ideally, this would be estimated by comparing the employment status of an individual with and without the training program at the same point in time.
- Counterfactuals are estimated using control groups.
- we can minimize the ethical concerns that may arise by denying treatment to the control group.
- Program activities produce outputs, and the resulting changes observed in the beneficiaries is the outcome. For example, in the case of vocational training, an outcome is employable skills.

In essence, the only design offered is a randomized control experimental design to test a model.

<http://siteresources.worldbank.org/INTCY/Resources/3957661186420121500/YDNVollI5Evaluation.pdf>

GOLD STANDARD:

METHODOLOGICAL APPROPRIATENESS



not

Methodological orthodoxy or rigidity

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

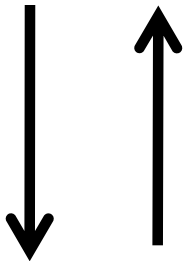
Fundamental Issue: How the World Is Changed

Top-down dissemination of

“proven models”

versus

**Bottoms-up adaptive
management**



Models vs. Principles

Identifying proven principles for
adaptive management
(bottoms-up approach)

versus

Identifying and disseminating
proven models
(top down approach)

Hypothesis

Much of the dynamics of
sustainability occurs in

the middle space

between the top and bottom,

where the top-down

and

bottoms-up forces

dynamically intersect and interact.

Conditions that challenge traditional model-testing evaluation

- **High innovation**
- **Development**
- **High uncertainty**
- **Dynamic**
- **Emergent**
- **Systems Change**



**Adaptive
Management**

Challenge:

**Situation
Recognition
and
Appropriate
Evaluation
Approach**

Conditions for which randomized control trials (RCTs) are especially appropriate

- A discrete, concrete intervention: singular, well- specified, high fidelity model
- Implementation can be standardized
- Implementation can be controlled
- Valid and reliable measures exist for the outcome to be tested
- Random assignment is possible
- Random assignment is ethical

Examples of Appropriate Use of RCTs

- Drug studies
- Fertilizer and crop yields
- A curriculum
- A single health practice:
 - brushing teeth
 - exercise regimen

Examples where RCTs are not appropriate:

- Complex, multi-dimensional and highly context-specific community interventions
- Dynamic innovations
- Adaptive management approaches –
no standardized model appropriate
- Ethical constraints
- Developmental, participatory,
beneficiary-directed, empowering
approaches

Classic RCT Problems

- Weak external validity
- Not double blind (the pharmaceutical gold standard)
- Controls rather than comparisons
- Poor treatment specification
- Misses important natural variations
- Greater within group variation than between group variation
- Experimental mortality
- Randomization bias
- Hinders project adaptability

Twelve methods that are appropriate and possible for evaluation of dynamic, innovative & responsive Youth Voluntary Service Programs

1. Cumulative, in-depth case studies of particular community interventions and outcomes (synthesis analysis)
2. Comparative pattern analysis of natural variation
3. Systems approaches & analysis

Diverse methods for impact evaluation

4. GEM (General elimination method)
or

Modus operandi approach, e.g.,
epidemiological methods

5. Theory-driven evaluation

6. *Realist evaluation*: What works for whom in what ways taking into account context? Focus on identifying and validating causal mechanisms.

Diverse methods for impact evaluation

7. Outcome mapping and contribution analysis (IDRC)

8. Beneficiary Assessment

(The World Bank, Lawrence Salmen and Eileen Kane)

9. Success Case Method

10. Comparative designs (as opposed to RCTs)

11. Regression discontinuity models

12. Interrupted time series analyses

Standards and criteria for “proof”

- Preponderance of evidence
- Cumulative evidence (synthesis of patterns of effectiveness across projects and countries)
- Triangulation (mixed methods)
- Utility
- Valid extrapolation of effective principles

CONTRASTS

Traditional evaluation...

- **Tests models**
- **Seeks best practices & “proven” models**

Complexity-based, Developmental Evaluation...

- **Supports innovation and adaptive management**
- **Seeks effective principles for adaptation**

Traditional Evaluation...

- **Renders definitive judgments of success or failure**

Developmental Evaluation...

- **Provides feedback, generates learning, supports direction or affirms changes in direction in real time for adaptation to new settings**

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**

Peer Review Network

- Create a network of youth practitioners and evaluators to provide peer review validation of proven principles and approaches using appropriate criteria and standards

References

Utilization-Focused Evaluation, 4th edition, 2008

Michael Quinn Patton, Sage Publications.

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

Getting to Maybe: How the World Is Changed? Frances Westley, Brenda Zimmerman, Michael Q. Patton, Random House Canada, 2006.