Evaluation design: a balancing act between feasibility, utility and scientific rigor

Jos Vaessen, PhD
Adviser evaluation methods, IEG

University of Vienna
June 24, 2019
Icebreaker

- Impact evaluation paradox
- Hierarchy-knowledge paradox
Outline

- Introduction
- Interpreting methodological rigor in evaluation
- Balancing rigor and utility
- Balancing rigor and feasibility
- Concluding remarks on the research-evaluation nexus
Introduction: research and evaluation

- Evaluation is applied social science research (Rossi)

- The rift between evaluation and the (social and behavioral) sciences
RESEARCH
Seek to generate new knowledge
Researcher-focused
Hypotheses
Make research recommendations
Publish results

METHODS & ANALYSIS

EVALUATION
Information for decision making
Stakeholder-focused
Key Questions
Recommendations based on key questions
Report to stakeholders

(LaVelle, 2010)
Utility

Rigor
Feasibility

Evaluation modality (evaluation theory)
Interpreting methodological rigor in evaluation

- Focus
- Consistency
- Reliability
- Validity
Rigor in valuing

- Factual versus value premises (Scriven)

- Evaluation focuses much more explicitly on the latter

- Valuing data and empirical findings against:
  - Intended objectives
  - Norms and standards
  - Institutional comparators
  - Past practice
  - What is realistically possible
  - What stakeholders need
  - …..
Balancing rigor and utility

- Different schools of thought in evaluation (‘theories of evaluation’)
- Different audiences, purposes and uses of evaluation
- The role of theory in evaluation: academic theory versus program theory
Research-based theory of PES

The logic of payments for environmental services. Source: Adapted from Pagiola and Platais (2007).
Program theory of PES

Multiple causal pathways and effects:

- Labor and employment market effects
- Biodiversity ‘corridor’ effects
- Environmental displacement effects
- National scaling-up effects
- (Inter) national replication effects
Balancing rigor and feasibility

- Finding out fast: scope versus depth trade-offs
- Finding out fast: the need for deductive and inductive inquiry
- Dealing with complexity
- Dealing with causality: an appeal to eclecticism
Scope versus depth

Synthetic ‘high-level’ ToC

Nested ‘detailed’ ToC
Figure A1. Methodological Design: Evaluation Components and their Relationship

Source: IEG
1. How did outreach evolve? Was there increased outreach among the rural poor?

2. What are the factors that explain outreach/access?

3. What are the implications for poverty alleviation?
Dimensions of COMPLEXITY

EMBEDDEDNESS AND THE NATURE OF THE SYSTEM

INTERVENTION

INSTITUTIONS AND STAKEHOLDERS

CAUSALITY AND CHANGE

Challenges in delimitation, sense-making, consensus-seeking, design implementation and use of EVALUATION

Source: Bamberger, Vaessen and Raimondo (2015)
Systems analysis (UR)

Institutional Behavior + Change in the Built Environment = Urban Flood Resilience

- Allocation and coordination of the cumulative efforts and their perceived success
- Institutional fragmentation and/or inconsistency
- Spatial implications of institutions for the future vulnerability of the built environment

World Bank Support
- Geographical location of current World Bank projects
- World Bank's contribution to change of institutional behavior

Recovery, Adapting, Transforming

Institutional Capacity:
- Financial capacity
  - Staff
  - Equipment
  - Regulatory reform
  - Guidance

World Bank projects' spatial effect in and outside the flood hazard zone

Financial and technical assistance [F]
- Training and capacity building for land administration [S]
- Institutional reform [S]
- Knowledge base and information system [S]

Ownership and tenure [C]
- Institutional coordination [L]
- Customary landowners [C]
- Inadequate staff and logistics [C]

Institutional crowdedness [C]
- Mitigation of flooding [C]

World Bank Input [I], Leverage points [L], Systems intervention [S], Context [C]
Positive relationship → Negative relationship ←
Evaluability challenge

WBG support (convening power / advisory / TA / funding) for:
- Institutional reform and collaboration
- Systems
- Processes
- Human resources
- Equipment
- Infrastructure
- …

Awareness, attitudes and behaviors of institutions, their capacities, and their (combined) interventions

Awareness, attitudes and behaviors of businesses, communities, citizens

Land uses and (changes) in the built environment
Dealing with causality: an appeal to eclecticism

<table>
<thead>
<tr>
<th>Overall impact question</th>
<th>Did the intervention make a difference?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific impact question</td>
<td>Did the intervention make a difference?</td>
</tr>
<tr>
<td></td>
<td>How much of a difference (on average)?</td>
</tr>
<tr>
<td></td>
<td>For whom? Under what circumstances?</td>
</tr>
<tr>
<td></td>
<td>How? Why so?</td>
</tr>
<tr>
<td>Causal question</td>
<td>Can we attribute the marginal (net) effect to the intervention? What is the net effect of other factors?</td>
</tr>
<tr>
<td></td>
<td>What role did the intervention play in producing the outcome?</td>
</tr>
<tr>
<td></td>
<td>What explains the outcome?</td>
</tr>
<tr>
<td>Causal theory</td>
<td>“Counterfactual”</td>
</tr>
<tr>
<td></td>
<td>“Multiple conjunctural”</td>
</tr>
<tr>
<td></td>
<td>“Generative” or “mechanism based”</td>
</tr>
<tr>
<td>Methods</td>
<td>e.g., (quasi) Experiments, stat modeling</td>
</tr>
<tr>
<td></td>
<td>e.g., Pattern-matching, QCA,…</td>
</tr>
<tr>
<td></td>
<td>e.g., Process tracing, in-depth case study,…</td>
</tr>
</tbody>
</table>

Source: Adapted from Befani, 2016 p. 20
Concluding remarks on the research-evaluation nexus

- Evaluation follows standards of methodological rigor in academic research to the extent ‘feasible and useful’

- In-depth rigorous causal analysis can be a stand-alone exercise or part of a larger evaluative exercise

- Existing academic research constitutes an important foundation for (part) of a larger evaluative exercise

- Academic research and evaluation overlap but evaluation has its own (trans)disciplinary characteristics
Thank you!
The data challenge