**Methodological Challenges
for Impact Evaluations of
Development Interventions**

Barbara Befani, PhD

03:06:2013

Contents

- The principles of the Paris Declaration
- Defining Impact
- Gold Standards of Evidence
- Impact questions
- Programme attributes
- Available Designs
- Mixed methods: what do we mix and why?
- Causal inference approaches
- Final remarks: implications of the principles of the PD

The Principles of the Paris Declaration: countries expected to the be in the lead

- Ownership
    - Country's own national development strategies
- Alignment
    - Donors' priorities aligned with the above
- Harmonisation
    - Donor coordination, avoiding duplication and minimizing transaction costs
- Management for Results
    - Aid must be proven to make a tangible difference in poor people's lives
- Mutual Accountability and Transparency
    - On use and results of aid
    - To citizens and parliaments

Defining Impact: three challenges

- Impact on what?
  - What is important? Value systems
- Defining success in conflict and uncertainty
  - avoiding failure
  - Achieving X = avoiding V, X and Z
- The attribution problem
  - Multitude of simultaneous interventions
  - Average impact?
    - With large variance, averages are not always useful!

Gold Standards of Evidence in Science (there is more than one!)

- RCTs? Yes, to generalize pilot results to a large, homogenous population, but also:
- Diagnostic Tests, to diagnose the presence of a condition, based on probability
  - Probability of Impact, given the observation of evidence
  - Probability of evidence, assuming there is impact
  - Sensitivity
  - Specificity
  - Predictive Power (the Bayes Formula)
- Bayesian Belief Nets (Bayes Nets)

Impact Questions

- Managing for results: of today or of tomorrow?
- Accountability vs. learning
  - Different types of learning at stake
- Measuring: how much?
  - Does it work? Did we do things right?
- Does it make a(ny) difference?
  - More open on the kind of difference...
- Assessing and understanding: how and why?
  - Why does it work? Did we do the right things? Under what circumstances does it work?
- Prediction and generalization
  - Will it work, elsewhere and elsewhen?
- For whom? (crosscutting)

Programme Attributes that pose methodological challenges

- Chaos, uncertainty, unpredictability
  - Conflict, natural disasters
- Indirect delivery
  - working through country governments, partnerships and local funds
  - Intermediate implementation steps
  - Local customisation
- Multiple delivery
  - Overlaps among donor activities
- Long-term and non linear change, limited predictability
- Multiple dimensions of change
  - Change that requires multiple changes
  - cultural change, attitudes and beliefs

Choosing and mixing methods

- Wide range of available designs
  - Substantial developments in social science methodology over the last 20 years, still not reflected in evaluators' toolkits
- Multiple approaches to causal inference
- Increasingly blurred distinction between quali, quanti
- Choice should be tailored to
  - Evaluation questions
  - Programme attributes

Range of available designs

- Table from SternEtAl

Blurred distinction between quali and quanti

- Quantitative data can be collected with participatory approaches (PADev, 'Who Counts?' J.Holland)
- Stress on validity rather than quali / quanti (G. Ton)
  - Construct, internal, external validity
- We don't always need real numbers (e.g. 7.432325)
  - Qualitative scales (very poor, poor, fair, good)
  - Dichotomous variables (yes / no)
  - Set-Theory vs. calculus

Multiple approaches to causal inference

- Multiple effects (intermediate outcomes, impact for whom?)
- Not only multiple effects, but also multiple causes
- Successionist: observation of simultaneous presence of cause and effect
- Generative: description of the causal process, of "inner workings" of causal mechanism
- Configurational: inner workings described as presence and absence of conditions to facilitate comparison and cross-case systematic synthesis
- More or less explicitly, all contemplate ways of building lists of possible causes and methods to successively eliminate them

The successionist approach to causality

- Causal Attribution, triggering change, sufficiency
- Simultaneous observation of cause and effect
- Single or multiple causality?
- In statistics, there are 'multiple' causes BUT causes are usually conceptualized as average contributions to an effect, can be added
- Causal contribution doesn't change depending on other causes, it is independent, not contextual
- The focus is on 'one cause-one effect' relationships
- Mill's Methods: agreement and difference
  – Counterfactual analysis is based on MoD
  – Their own logic of causal elimination

The Generative Approach to Causality

- Causal contribution, preparing the ground for change, necessity
- Not merely about simultaneous observation of one cause and one effect
- Conceptualizing the cause as a potentially complex mechanism / process
- Describing the transformation from cause to effect
  - The process whereby the cause transforms into the effect
  - the inner workings of the causal mechanism
- Not merely "attribution", but more like "explanation"
- Other alternative causes / explanation / processes are "eliminated" using more or less explicit probability principles

The configurational approach to causality 1

- Inbetween successionism and generativism
- Both contribution and attribution, necessity and sufficiency, causes that prepare the ground and causes that trigger change
- INUS and SUIN causes
  - INUS: trigger change, sufficient for change but only within a specific context (a given combination of necessary factors)
  - SUIN: one of a group of equally necessary, ground-preparing causes (either is sufficient to meet the requirements)

The configurational approach to causality 2

- Multiple observations of simultaneous presence of cause(s) and effect(s)
- But observations must be comparable
- The description of the causal process / mechanism is highly standardized
    - Presence / absence of conditions (dichotomous)
    - Membership scores to fuzzy sets (of combinations of conditions)
- Causes are eliminated by the one-difference rule
- "the best of both worlds":
    - Both explanation and generalization
        - Although a less sophisticated explanation
        - And a more limited generalization

Mixing methods: WHAT do we mix and how?

- What do we mix? A broad range of:
  - Criteria to choose most relevant change, depending also on who we involve
    - What impact for whom?
  - Causal inference models
    - Who builds the list of possible causes?
      - Availability bias
    - Attribution or contribution?
    - Explanation or generalization?
  - Designs
  - Methods
  - Data collection techniques

Mixing methods: what do we mix and HOW? 1

- How do we mix? Depending on:
  - Impact questions:
    - Participatory approaches are required to answer 'what impact for whom' and addressing causality
    - Did it work? Did we do things right? Most likely require successionist approaches to causality, regression-based and (quasi) experimental designs
    - How and why did it work (or not)? Did we do the right things? Most likely require generative approaches to causality, indepth "studies of the case", process tracing, impact pathways, etc.

Mixing methods: what do we mix and HOW? 2

- How do we mix? Depending on:
  - Impact questions:
    - Did it make a difference? Under what circumstances did it work better? Will it work elsewhere / elsewhen? Most likely requires configurational approaches to causality, systematic cross-case comparison, synthesis
  - Programme attributes:
    - Conflict, natural disasters
      - Chaos theory, probability, implementation proc.s capable of handling uncertainty & emergence, RTE

Mixing methods: what do we mix and HOW? 2

- How do we mix? Depending on programme attributes:
- Change that requires multiple changes, "wicked" problems with positive and negative, unpredictable feedback loops
  - Systems thinking, holistic approach, generative c, agent-based modeling
- Long term and non linear change
  - Process tracing, systems dynamics, non linear modeling, generative c
- Indirect Delivery, local customization
  - Configurational c, studies of the case
- Multiple delivery
  - Configurational c, studies of the case
  - Joint, multi-donor evaluations

Conclusion: some implications of the PD principles

- Ownership
  - What impact for whom?
- Alignment
  - Government is an important stakeholder of evaluations (joint evaluations)
- Harmonisation
  - Multi-donor evaluations, holistic app, case studies
- Management for Results
  - Evidence-based policy, evaluation research
- Mutual Accountability and Transparency
  - Causal attribution and contribution
  - Rigour in choice and application of methods