3ie-LIDC Seminar: What works in international development

Attribution of cause and effect in small n impact evaluations

www.3ieimpact.org
www.lidc.org.uk
SMALL n IMPACT EVALUATION

Howard White and Daniel Phillips

3ie
Impact evaluations answer the question as to what extent the intervention being evaluated altered the state of the world. All project-affected persons (PAPs) and all outcomes and types of outcome, intended and unintended, are considered in these evaluations. Outputs or outcomes are also part of the analysis.
**Definitions II**

- What is small n?
  - Data on too few units of assignment to permit tests of statistical significance between treatment and a comparison group

- Why do you have small n?
  - Small N
  - Heterogeneity
  - Budget
  - NOT
    - Universal or complex interventions
**Definitions III**

- Small n impact evaluation is
  - Still about attribution, i.e. what difference did the intervention make?
  - And so outcome monitoring alone is not enough
  - About demonstrating a case ‘beyond reasonable doubt’ of the link between intervention and change in the state of the work, which is more than ‘simple association’, e.g. policy reform
  - Includes tricky cases of declining counterfactual

- Difference between small and large n
  - Large n establishes causation through statistical means
  - Small n builds up case based on weight of evidence, strength of argument, and absence of other plausible explanations
  - In mixed methods designs the large n component is the DNA evidence that can usually clinch the causal argument
The counterfactual

Outcome

Time

Factual

Counterfactual
We would have done it anyway
THE COUNTERFACTUAL: WOULD HAVE HAPPENED ANYWAY

Outcome

Counterfactual

Factual

Time
THE COUNTERFACTUAL II
What is the counterfactual?
DEFINITIONS IV
CONTRIBUTION VERSUS ATTRIBUTION

- Impact evaluation is defined as attribution
- Possible cases for wanting to use contribution instead:
  - Multiple factors – but IE is meant to disentangle them, and attribution doesn’t mean sole attribution
  - Complementarities (two necessary conditions, and neither one sufficient, e.g. school feeding) – then state what they are
  - Over-determination (two sufficient conditions, both present) – determine which is most cost effective e.g. WSS
  - Complexity – then the black box may be a useful approach (remember Semmelweis)
- So contribution, which means same as attribution, because attribution analysis in IE seem to be limited, but these presumed limitations are mostly absent
When school feeding works

Understand context to look at sources of heterogeneity
HYGIENE AND SANITATION: SUBSTITUTES OR COMPLEMENTS?

Need a factorial design
APPROACHES TO SMALL n IMPACT EVALUATION

- **Explanatory approaches**
  - Contribution analysis, general elimination methodology, realist evaluation, process tracing
  - Based around theory of change (causal chain)
  - Explicit accounting for context and ‘other factors’
  - Possibly generate testable hypotheses, to be tested using mixed methods

- **Participatory approaches**
  - Method for Impact Assessment of Programs and Projects (MAPP), most significant change, success case method, outcome mapping
  - Uses participatory data to analyze cause and effect (role programme has played in changes at community and individual level)
SO WHERE DOES THIS LEAVE US?

- **Common elements**
  - Clear statement of intervention
  - Lay out theory of change, allowing for context and external factors (e.g. CCTV)
  - Document each link in causal chain
  - If things happened as planned, and intended outcome observed, then conclude causation (or not e.g. Peru microcredit) – using triangulation and other evidence e.g. Ghana hospital

- **But the last step is getting a bit dodgy. Something is missing – what constitutes valid evidence of a link in the causal chain?**
WHAT IS VALID CAUSAL EVIDENCE?

- Want an approach which uncovers the ‘true causal relationship’ in the study population (internal validity)
- In large n studies threats to internal validity can come from sampling error and selection bias
- Analogously, in small n studies bias arises if there is a systematic tendency to over or under-estimate the strength of the causal relationship
IS THERE SYSTEMATIC BIAS IN QUALITATIVE DATA COLLECTION AND ANALYSIS?

“There exist significant sources of bias in both the collection and analysis of qualitative data”

“These biases arise both in the responses given and the way in which evaluators interpret these data”

“These biases are likely to result in the systematic over-estimate of programme impact”
COURTESY

BIAS

IF YOU CAN'T SAY ANYTHING NICE, DON'T SAY ANYTHING AT ALL.

THAT'S GOING TO BE REALLY DIFFICULT.
PERSPECTIVE DEPENDS ON WHERE YOU ARE STANDING
Of course I have a grandiose sense of self-importance. Who doesn’t?
The Barry Manilow t-shirt experiment
EXAMPLES

- Dating of policy reform
- Fundamental error of attribution: Missing actors or context, e.g. local policy environment
- People we don’t speak to (trade unions, parliamentarians .....
- Generally over-stating role of intervention e.g. DANIDA livelihoods

- The way in which evaluators interpret data reinforces the bias of respondents to overstate their role
SIMILAR PERSON AND EXPOSURE BIAS

Blah, blah, blah, blah....

That’s not how it was at all
WHAT TO DO

Well structured research with systematic analysis of qualitative data
Define intervention to be evaluated

Theory of change

Identify mix of methods to answer each question

Identify evaluation questions

Data collection and analysis plan