IMPACT EVALUATION OF EDUCATION INTERVENTIONS

ALL CHILDREN READING ASIA WORKSHOP
BANGKOK, SEPTEMBER 20TH 2012

Richard L. Sawyer, IPA, Country Director, Mongolia
Correlation vs. Causation

I USED TO THINK
CORRELATION IMPLIED
CAUSATION.

THEN I TOOK A
STATISTICS CLASS.
NOW I DON'T.

SOUNDS LIKE THE
CLASS HELPED.
WELL, MAYBE.
What is the most cost-effective intervention to increase attendance in school?

- Information on returns to schooling
- School based de-worming
- School meals
- School uniforms
- Girls merit scholarships
- Conditional cash transfers
Surprisingly little hard evidence on what works

Can do more with given budget with better evidence

If people and organization knew money was going to programs that worked, it could help increase pot for anti-poverty programs

Instead of asking “do aid/development programs work?” should be asking:
- Which work best, why and when?
- How can we scale up what works?
What goes into policymaking?

Knowledge:
• Evidence
• Experience

Ideology

Resources

Policies & Projects
Expand the Role of Evidence
What Makes a Good Evaluation?

- Ask the right questions
  - For accountability
  - For lesson learning
- Answer those questions in unbiased and definitive way
- Share answers regardless of the results
- To do that you need a model:
  - Logical framework/theory of change
  - Who is the target?
  - What are their needs?
  - What is the program seeking to change?
  - What is the precise program or part of program being evaluated?
Logical Framework

Needs

Needs assessment

Program evaluation

Input

Process evaluation

Output

Impact evaluation

Outcome

Impact (primary outcome)

Goal
Role of Impact Evaluation

- The program happened, how did it change lives?
- What does Theory of Change say we might expect to change?
- Primary outcome (impact)
- Also distributional questions
How to Measure Impact?

- What would have happened in the absence of the program?

- Take the difference between

  What happened (with the program)
  - What would have happened (without the program)
  = IMPACT of the program
How to Measure Impact?

- Time
- Primary Outcome
- Impact
- Counterfactual

Diagram showing time on the x-axis and primary outcome on the y-axis. The intervention is marked with a red vertical line, and the impact is shown as an upward arrow. The counterfactual line is depicted as a green line.
How to Measure Impact?

- Time
- Primary Outcome
- Impact
- Counterfactual
- Intervention

Diagram showing the relationship between time, primary outcome, intervention, and impact with a counterfactual.
How to Measure Impact?

What is it?

Primary Outcome

Impact

Counterfactual

Time

Intervention

Primary Outcome

Time
How Impact Differs from Process?

- When we answer a process question, we need to describe what happened.

- When we answer an impact question, we need to compare what happened to what would have happened without the program.
The counterfactual represents how program participants would have performed in the absence of the program.

**Problem:** Counterfactual cannot be observed

**Solution:** We need to “mimic” or construct the counterfactual.

Different impact evaluation methodologies differ in how they construct the counterfactual.
Methodologies for Impact Evaluation

- Non-experimental
  - Regression
  - Pre-post
  - Simple Comparison

- Quasi-experimental
  - Instrumental Variables
  - Regression Discontinuity Design

- Experimental
  - Randomized Evaluations

→ Key Question: How good is the counterfactual?
Randomized Evaluations

Time

Primary Outcome

Impact

Counterfactual

Intervention

Counterfactual
Pre-Post

Primary Outcome vs Time

Intervention

Impact
Inappropriate Comparison Group

- **Primary Outcome**
  - **Time**
  - **Intervention**
  - **Impact**
  - **Comparison Group**
  - **Counterfactual**

**Graph Diagram:**
- Y-axis: Primary Outcome
- X-axis: Time
- Graph shows trends over time with different colored lines representing different groups and impacts before and after an intervention.
Ex Post Studies … we know too little
Problem: Reading ability of many students was well below grade level. Salary structure in Philippines left little money for training, textbooks or buildings

Question: Initial reading program was not enough, can a month-long reading marathon have more substantial effects?

Design: 31-day read-a-thon: daily reading activities (storytelling sessions, reading games, and posters), reading charts, diaries with P4 students
- Baseline, 2 follow up surveys
- Reading skills, knowledge of subjects
Encouraging increased use of age appropriate reading materials by students is a viable strategy for improving student’s reading skills.

- Immediately after program, students report that the number of books they read at school increased from 1.9 to 4.2 in the last week and from 2.3 to 9.5 in the last month.
- Reading scores increased by 0.13 standard deviations.

Three months later:

- Positive effects persisted - treatment students still read 3.1 books more in the previous month than the comparison group.
- Students in the treatment group also had reading scores 0.06 standard deviations higher than the comparison students
Problem: A 2005 survey found that 44% of Indian children age 7 to 12 could not read a basic paragraph and 50% could not do simple subtraction even though most are enrolled in school. Even in urban India, the learning levels are very low.

Question: Can community teachers (balsakhis) improve learning outcomes for low performing students simple by pulling them out of classes and working with them for 2 hours a week?

Design: In the 2001 school year approximately half of the schools were given a tutor for grade 3, and the other half were given a tutor for grade 4 - which school received which was randomized. In 2002, the schools were given a tutor for the previously untreated grade.

Researchers: Abhijit Banerjee
            Esther Duflo
            Leigh Linden

Partners: Pratham
          World Bank

Where: India

When: 2001

Sample: 122 primary schools in Vadodara, 77 primary schools in Mumbai, India
The number of students in the bottom third of program classes who passed basic competency tests increased by nearly 8%.

Those in the top third who passed increased by 4%.

Only 2-3% of children in the bottom third could do long division at the start of the year, but by the end of the year 40% could in program classes compared to 28% in non-treatment classes.
The Method Matters!

Estimates for Balsakhi Program

<table>
<thead>
<tr>
<th>Method</th>
<th>Impact Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Pre-post</td>
<td>26.42*</td>
</tr>
<tr>
<td>(2) Simple Difference</td>
<td>-5.05*</td>
</tr>
<tr>
<td>(3) Difference-in-Difference</td>
<td>6.82*</td>
</tr>
<tr>
<td>(4) Regression</td>
<td>1.92</td>
</tr>
<tr>
<td>(5) Randomized Experiment</td>
<td>5.87*</td>
</tr>
</tbody>
</table>

*: Statistically significant at the 5% level
Developing an Evaluation Strategy

- Things to Keep in Mind:
  - Key questions can be answered from impact evaluations, process evaluations tell you what you are measuring
  - A few high quality impact studies are worth more than many poor quality ones
  - If you ask the right question, your organization will be more likely to use the knowledge gained
Where to Start?

- **Policy relevance** – Based off of what is already known, what are the most important policy questions? Which programs would be mostly like to need cost effectiveness information in order acquire funding and approval?

- **Policy Choices** – What variations are you trading off between different programs?

- **Operational Tests** – How will the program be implemented? How will you target the program? How will you monitor it? What will you need to do to ensure usage and take up?

Often the best designs are simple and focused.
Where to Start?

- What are the key questions for your organization?
  - How many of these can you already answer from existing rigorous impact evaluations? Improved process evaluations?

- Select top priority questions for impact evaluation

- Establish a plan for answering them
  - Take advantage of roll out of new programs or set up experimental sites
  - Take time to carefully map out the logic of your program and think about how outputs and outcomes can be measured
Answer to the Question

J-PAL COST EFFECTIVENESS: additional years of student attendance per $100 spent

www.povertyactionlab.org

19.8 years  
6.1 years § 
2.9 years § 
0.72 years § 
0.27 years § 
0.032 year § 
0.047 year §

Information  
School-Based Deworming KENYA  
Iron Tablets & Deworming INDIA  
Subsidized School Uniforms KENYA  
Merit Scholarships for Girls KENYA  
Primary School CCTs (w/ transfers) MEXICO  
Primary School CCTs (w/out transfers) MEXICO

1  
2  
3  
4  
5  
6
More Information?

IPA’s research findings:
http://www.poverty-action.org/project-evaluations/search

J-PAL’s research findings
http://www.povertyactionlab.org/search/apachesolr_search?filters=type:evaluation

The ‘ladybird book about RCTs’:
‘Test, Learn, Adapt’, Cabinet Office, Ben Goldacre et al.
A Few Good Books

More Than Good Intentions

Abhijit V. Banerjee and Esther Duflo
Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty

It Ain’t What You Give: It’s the Way That You Give It
Caroline Fiennes
THANK YOU!

www.poverty-action.org

rsawyer@poverty-action.org