What is Evaluation?

Marc Shotland
J-PAL Global
Course Overview

1. What is evaluation?
2. Measuring impacts (outcomes, indicators)
3. Why randomize?
4. How to randomize?
5. Sampling and sample size
6. Threats and Analysis
7. RCT: Start to Finish
8. Cost Effectiveness Analysis and Scaling Up
What is Evaluation?

Evaluation
Program Evaluation
Impact Evaluation
Program Evaluation

Evaluation

Program Evaluation

Impact Evaluation
Monitoring and Evaluation

- Evaluation
- Program Evaluation
- Impact Evaluation

Monitoring
What’s the difference between:
Monitoring and Evaluation

A. Nothing. They are different words to describe the same activity
B. Monitoring is conducted internally, Evaluation is conducted externally
C. Monitoring is for management, Evaluation is for accountability
D. Don’t know
E. Other
Program Evaluation

Monitoring

Evaluation

Program Evaluation

Impact Evaluation
Components of Program Evaluation

- Needs Assessment
- Program Theory Assessment
- Process Evaluation
- Impact Evaluation
- Cost Effectiveness
- What is the problem?
- How, in theory, does the program fix the problem?
- Does the program work as planned?
- Were its goals achieved? The magnitude?
- Given magnitude and cost, how does it compare to alternatives?
Evaluation should usually be conducted:

A. Externally and independent from the implementers of the program being evaluated
B. Externally and closely integrated with program implementers
C. Internally
D. Don’t know
Who is this evaluation for?

- Academics
- Donors
  - Their Constituents
- Politicians / policymakers
- Technocrats
- Implementers
- Proponents, Skeptics
- Beneficiaries
Who is your *most important audience* for evaluation?

A. Agency leadership
B. Donor Politicians / policymakers
C. Donor Constituents
D. Academics
E. Indonesian Politicians / policymakers
F. Technocrats
G. Implementers
H. Proponents, Skeptics
I. Beneficiaries
Does Aid Work?
“I have identified the specific investments that are needed [to end poverty]; found ways to plan and implement them; [and] shown that they can be affordable.”

Jeffrey Sachs
End of Poverty
Aid Pessimists

“After $2.3 trillion over 5 decades, why are the desperate needs of the world's poor still so tragically unmet?

Isn't it finally time for an end to the impunity of foreign aid?”

Bill Easterly
The White Man’s Burden
Where does J-PAL fall in this debate?

- Somewhere in the middle
- We believe aid *can* help
- Too often it does not
- We don’t know when it does and when it doesn’t
- So we have a lot to learn about *how* Aid can help
How can impact evaluation help us?

- Surprisingly little hard evidence on what works
- Can do more with given budget with better evidence
- If people knew money was going to programs that worked, could help increase pot for anti-poverty programs
- Instead of asking “do aid/development programs work?” should be asking:
  - Which programs work best, why and when?
  - Which concepts work, why and when?
  - How can we scale up what works?
- Add to our body of evidence
  - part of a well-thought out evaluation (research) strategy
Programs and their Evaluations: where do we start?

Intervention
- Start with a problem
- Verify that the problem actually exists
- Generate a theory of why the problem exists
- Design the program
- Think about whether the solution is cost effective

Program Evaluation
- Start with a question
- Verify the question hasn’t been answered
- State a hypothesis
- Design the evaluation
- Determine whether the value of the answer is worth the cost of the evaluation
What do you think is the most cost-effective way to reduce diarrhea?

A. Develop piped water infrastructure
B. Improve existing water sources
C. Increase supply of and demand for chlorine
D. Education on sanitation and health
E. Improved cooking stoves for boiling water
F. Improve sanitation infrastructure
Identifying the problem

NEEDS ASSESSMENT
The Need

- Nearly 2 million children die each year from diarrhea
- 20% all child deaths (under 5 years old) are from diarrhea
The Likely Problem

• 13% of world population lacks access to “improved water sources”
The Goal

• MDG: “reduce by half the proportion of people without access to sustainable drinking water”
The Solution(s)
Really the Problem?

- *Quantity* of water is a better determinant of health than *quality* of water (Curtis et al, 2000)
- Water quality helps little without hygiene (Esrey, 1996)
  - 42% live without a toilet at home
- Nearly 2.6 billion people lack any improved sanitation facilities ([WHO](https://www.who.int/)
- People are more willing to pay for convenient water than clean water
- Chlorine is very cheap,
  - In Zambia, $0.18 per month for a family of six
  - In Kenya, $0.30 per month
- Yet less than 10% of households purchase treatment
- 25% of households reported boiling their drinking water the prior day

Alternative Solution(s)?
Devising a Solution

• What is the theory behind your solution?
• How does that map to your theory of the problem?
Blueprint for Change

PROGRAM THEORY ASSESSMENT
Program Theory Assessment

- Logical Framework (LogFrame, LFA)
- Theory of Change
- Results Framework
- Outcome Mapping

- Causal chain
- Causal model
- Cause map
- Impact pathways
- Intervention theory
- Intervention framework
- Intervention logic
- Investment logic
- Logic model
- Outcomes chain
- Outcomes hierarchy
- Outcome line
- Program logic
- Program theory
- Programme theory
- Results chain
- Theory-based evaluation
- Theory-driven evaluation
- Theory-of-action

Source: Patricia Rogers
Contaminated water is primary source of illness

Less Diarrhea

Drink Clean water

Have access to clean water at home

Choose to collect only clean water

Understand benefits of clean water

Choose to drink only clean water

Know which water is clean

Access to clean water at source

No recontamination

Hygiene practices

Clean method of extracting water

Sufficient water
## Log Frame

### Impact (Goal/ Overall objective)

<table>
<thead>
<tr>
<th>Objectives Hierarchy</th>
<th>Indicators</th>
<th>Sources of Verification</th>
<th>Assumptions / Threats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower rates of diarrhea</td>
<td>Rates of diarrhea</td>
<td>Household survey</td>
<td>Waterborne disease is primary cause of diarrhea</td>
</tr>
</tbody>
</table>

### Outcome (Project Objective)

<table>
<thead>
<tr>
<th>Outputs</th>
<th>Inputs (Activities)</th>
<th>Source water is cleaner; Families collect cleaner water</th>
<th>Source protection is built</th>
<th>Protection is present, functional</th>
<th>Source visits/surveys</th>
<th>Sufficient materials, funding, manpower</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household drink cleaner water</td>
<td>Source water is cleaner; Families collect cleaner water</td>
<td>Source protection is built</td>
<td>Protection is present, functional</td>
<td>Source visits/surveys</td>
<td>Sufficient materials, funding, manpower</td>
<td></td>
</tr>
</tbody>
</table>

### Adjustments

- (Δ in) drinking water source; E. coli CFU/100ml

### Notes

Program Theory Assessment

• How will the program address the needs put forth in your needs assessment?
  – What are the prerequisites to meet the needs?
  – How and why are those requirements currently lacking or failing?
  – How does the program intend to target or circumvent shortcomings?
  – What services will be offered?
Making the program work

PROCESS EVALUATION
Process Evaluation

• Supply Side
  – Logistics
  – Management
• Demand Side
  – Assumption of knowledge, preferences
  – Assumptions of response
Process Evaluation: Logistics

• Construction
  – Construct spring protection
  – Installing fencing
  – Installing drainage

• Maintenance
  – Patch concrete
  – Clean catchment area
  – Clear drainage ditches
Process Evaluation: Supply Logistics
Monitoring and Evaluation

- Monitoring
- Impact Evaluation
- Program Evaluation
- Evaluation
Process Evaluation: Demand-side

- Do households collect water from improved source?
- Does storage become re-contaminated?
- Do people drink from “clean” water?
Process was okay, so....

• What happened to diarrhea?
Measuring how well it worked

IMPACT EVALUATION
Did we achieve our goals?

• Primary outcome (impact): did spring protection reduce diarrhea?

• Also distributional questions: what was the impact for households with good v. bad sanitation practices?
What is Impact?

Impact is the change in the primary outcome caused by the intervention. It is measured by comparing the outcome after the intervention to what it would have been in the absence of the intervention (the counterfactual).
How to measure impact?

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

• Take the difference between
  what happened (with the program) ...and
  - what would have happened (without the program)
  = IMPACT of the program
Constructing the Counterfactual

• Counterfactual is often constructed by selecting a group not affected by the program

• Randomized:
  – Use random assignment of the program to create a control group which mimics the counterfactual.

• Non-randomized:
  – Argue that a certain excluded group mimics the counterfactual.
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 “gold standard” for Impact Evaluation

RANDOMIZED EVALUATION
Randomly sample from area of interest
Randomly **sample** from area of interest

Randomly **assign** to treatment and control

Randomly **sample** from both treatment and control
**Spring Cleaning Sample**

- **Total Population**: (562 springs)
- **Target Population**: (200)
- **Not in evaluation**: (0)
- **Evaluation Sample**: (200)
- **Random Assignment**:
  - Year 1: (50)
  - Year 2: (50)
  - Years 3,4: (100)
Impact

• 66% reduction in source water E coli concentration
• 24% reduction in household E coli concentration
• 25% reduction in incidence of diarrhea
<table>
<thead>
<tr>
<th>Intervention</th>
<th>Impact on Diarrhea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring protection (Kenya)</td>
<td>25% reduction in diarrhea incidence for ages 0-3</td>
</tr>
</tbody>
</table>
## Making Policy from Evidence

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Impact on Diarrhea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring protection (Kenya)</td>
<td>25% reduction in diarrhea incidence for ages 0-3</td>
</tr>
<tr>
<td>Source chlorine dispensers (Kenya)</td>
<td>20-40% reduction in diarrhea</td>
</tr>
<tr>
<td>Home chlorine distribution (Kenya)</td>
<td>20-40% reduction in diarrhea</td>
</tr>
<tr>
<td>Hand-washing (Pakistan)</td>
<td>53% drop in diarrhea incidence for children under 15 years old</td>
</tr>
<tr>
<td>Piped water in (Urban Morocco)</td>
<td>0.27 fewer days of diarrhea per child per week</td>
</tr>
</tbody>
</table>
Evidence-Based Policymaking

COST-EFFECTIVENESS ANALYSIS
Cost-Effectiveness Diagram

Cost-Effectiveness: Diarrheal Incidents Avoided per $1000
Sensitivity to Population Density

- Free Chlorine Dispensers at Water Sources
  - Kenya
- Free Delivery of Chlorine
  - Kenya
- Encasing Water Sources in Concrete
  - Kenya
- Free Delivery of Chlorine
  - Pakistan
- Handwashing Promotion with Free Soap
  - Pakistan

Legend:
- Source improvements
- Changing behavior
- Chlorine treatment
When is a good time to do a randomized evaluation?

A. After the program has begun and you are not expanding it elsewhere
B. When a positive impact has been proven using rigorous methodology
C. When you are rolling out a program with the intention of taking it to scale
D. When a program is on a very small scale e.g one village with treatment and one without
When to do a randomized evaluation?

- When there is an important question you want/need to know the answer to
- Timing--not too early and not too late
- Program is representative not gold plated
  - Or tests an basic concept you need tested
- Time, expertise, and money to do it right
- Develop an evaluation plan to prioritize
When NOT to do an RE

• When the program is premature and still requires considerable “tinkering” to work well

• When the project is on too small a scale to randomize into two “representative groups”

• If a positive impact has been proven using rigorous methodology and resources are sufficient to cover everyone

• After the program has already begun and you are not expanding elsewhere
Developing an evaluation strategy

- Start with a question
- Verify the question hasn’t been answered
- State a hypothesis
- Design the evaluation
- Determine whether the value of the answer is worth the cost of the evaluation

- With key questions answered from impact evaluations, process evaluation can give your overall impact
- A few high quality impact studies are worth more than many poor quality ones

- If you ask the right question, you’re more likely to care
Components of Program Evaluation

- Needs Assessment
- Program Theory Assessment
- Process Evaluation
- Impact Evaluation
- Cost Effectiveness

- What is the problem?
- How, in theory, does the program fix the problem?
- Does the program work as planned?
- Were its goals achieved? The magnitude?
- Given magnitude and cost, how does it compare to alternatives?