What is Evaluation?

Adam Osman
J-PAL
Course Overview

1. What is Evaluation?
2. Outcomes, Impact, and Indicators
3. Why Randomize
4. How to Randomize
5. Threats and Analysis
6. Sampling and Sample Size
7. Project from Start to Finish
8. Generalizability
What is Evaluation?

- Evaluation
- Program Evaluation
- Impact Evaluation
Program Evaluation
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
## 5 Components of Program Evaluation

<table>
<thead>
<tr>
<th>1. Needs Assessment</th>
<th>A. What is the problem?</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Program Theory Assessment</td>
<td>B. How, in theory, does the program fix the problem?</td>
</tr>
<tr>
<td>3. Process Evaluation</td>
<td>C. Does the program work as planned?</td>
</tr>
<tr>
<td>4. Impact Evaluation</td>
<td>D. Were its goals achieved? The magnitude?</td>
</tr>
<tr>
<td>5. Cost Effectiveness</td>
<td>E. Given magnitude and cost, how does it compare to alternatives?</td>
</tr>
</tbody>
</table>
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?

A. Politicians / policymakers
B. Constituents
C. Donors
D. Donor Politicians / policymakers/ constituents
E. Academics
F. Technocrats / Experts/ Think Tanks
G. Implementers
H. Proponents, Skeptics
I. Beneficiaries
Who is your *most important audience* for evaluation?

A. Politicians / policymakers
B. Constituents
C. Donor leadership
D. Donor politicians / policymakers/ constituents
E. Academics
F. Technocrats / Experts/ Think Tanks
G. Implementers
H. Proponents, Skeptics
I. Beneficiaries
Programs and their Evaluations: where do we start?

**Intervention**
A. Start with a problem
B. Verify that the problem actually exists
C. Generate a theory of why the problem exists
D. Design the program
E. Think about whether the solution is cost effective

**Program Evaluation**
A. Start with a question
B. Verify the question hasn’t been answered
C. State a hypothesis
D. Design the evaluation
E. Determine whether the value of the answer is worth the cost of the evaluation
An Example

WATER, SANITATION & HEALTH
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

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

A. Bad Water

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

A. MDG: “reduce by half the proportion of people without access to sustainable drinking water”
The Solution(s)
A. Water quality helps little without hygiene (Esrey, 1996)
   A. 42% live without a toilet at home
B. Nearly 2.6 billion people lack any improved sanitation facilities (WHO)
C. Quantity of water is a better determinant of health than quality of water (Curtis et al, 2000)
D. People are more willing to pay for convenient water than clean water
E. Chlorine is very cheap,
   A. In Zambia, $0.18 per month for a family of six
   B. In Kenya, $0.30 per month
F. Yet less than 10% of households purchase treatment
Alternative Solution(s)?
Devising a Solution

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

PROGRAM THEORY

ASSESSMENT
Program Theory Assessment

A. Logical Framework (Log Frame)
B. Theory of Change
C. Results Framework
D. Outcome Mapping

A. Causal chain
B. Causal model
C. Cause map
D. Impact pathways
E. Intervention theory
F. Intervention framework
G. Intervention logic
H. Investment logic
I. Logic model
J. Outcomes chain
K. Outcomes hierarchy
L. Outcome line
M. Program logic
N. Program theory
O. Programme theory
P. Results chain
Q. Theory-based evaluation
R. Theory-driven evaluation
S. Theory-of-action

Source: Patricia Rogers
## Log Frame

<table>
<thead>
<tr>
<th>Objectives Hierarchy</th>
<th>Indicators</th>
<th>Sources of Verification</th>
<th>Assumptions / Threats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact (Goal/ Overall objective)</td>
<td>Lower rates of diarrhea</td>
<td>Rates of diarrhea</td>
<td>Waterborne disease is primary cause of diarrhea</td>
</tr>
<tr>
<td></td>
<td>Household survey</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcome (Project Objective)</td>
<td>Households drink cleaner water</td>
<td>(Δ in) drinking water source; E. coli CFU/100ml</td>
<td>Shift away from dirty sources. No recontamination</td>
</tr>
<tr>
<td></td>
<td>Household survey, water quality test at home storage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outputs</td>
<td>Source water is cleaner; Families collect cleaner water</td>
<td>E. coli CFU/100ml;</td>
<td>continued maintenance, knowledge of maintenance practices</td>
</tr>
<tr>
<td></td>
<td>Water quality test at source</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inputs (Activities)</td>
<td>Source protection is built</td>
<td>Protection is present, functional</td>
<td>Sufficient materials, funding, manpower</td>
</tr>
<tr>
<td></td>
<td>Source visits/ surveys</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Program Theory Assessment

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

PROCESS EVALUATION
With Process Evaluation

A. Was the program implemented as planned
B. Did people respond as expected
C. If it were…
   A. What about the concept?
Process Evaluation

A. Supply Side
   A. Logistics
   B. Management

B. Demand Side
   A. Assumption of knowledge, preferences
   B. Assumptions of response
Process Evaluation: Logistics

A. Construction
   A. Construct spring protection
   B. Installing fencing
   C. Installing drainage

B. Maintenance
   A. Patch concrete
   B. Clean catchment area
   C. Clear drainage ditches
Process Evaluation: Supply Logistics
A. Do households collect water from improved source?
B. Does storage become re-contaminated?
C. Do people drink from “clean” water?
Measuring how well it worked

IMPACT EVALUATION
Did we achieve our goals?

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

B. Also distributional questions: what was the impact for households with good v. bad sanitation practices?
What is the impact of this program?

- Primary Outcome

- Program starts

- Time
What is the impact of this program?

1. Positive
2. Negative
3. Zero
4. Not enough info
What is the impact of this program?

1. Positive
2. Negative
3. Zero
4. Not enough info
How to measure impact?

*Impact* is defined as a comparison between:

1. the outcome some time after the program has been introduced

2. the outcome at that same point in time had the program not been introduced (the “counterfactual”)

The *counterfactual* represents the state of the world that program participants would have experienced in the absence of the program (i.e. had they not participated in the program)

*Problem:* Counterfactual cannot be observed

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

• Usually done by selecting a group of individuals that did not participate in the program

• This group is usually referred to as the control group or comparison group

• How this group is selected is a key decision in the design of any impact evaluation
Selecting the comparison group

• Idea: Select a group that is exactly like the group of participants in all ways except one: their exposure to the program being evaluated

• Goal: To be able to attribute differences in outcomes between the group of participants and the comparison group to the program (and not to other factors)
How to measure impact?

A. What would have happened in the absence of the program?

B. Take the difference between
   
   what happened (with the program) …and
   
   - what would have happened (without the program)
   
   = IMPACT of the program
What is the impact of this program?
Impact: What is it?

Program starts

Counterfactual

Impact

Time

Primary Outcome
Impact: What is it?

Time

Program starts

Primary Outcome

Counterfactual

Impact
Impact evaluation methods

1. Randomized Experiments
   • Also known as:
     – Random Assignment Studies
     – Randomized Field Trials
     – Social Experiments
     – Randomized Controlled Trials (RCTs)
     – Randomized Controlled Experiments
Impact evaluation methods

2. Non- or Quasi-Experimental Methods
   a. Pre-Post
   b. Simple Difference
   c. Differences-in-Differences
   d. Multivariate Regression
   e. Statistical Matching
   f. Interrupted Time Series
   g. Instrumental Variables
   h. Regression Discontinuity
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
Impact

A. 66% reduction in source water E coli concentration
B. 24% reduction in household E coli concentration
C. 25% reduction in incidence of diarrhea
## 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>
</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

- Source improvements
- Changing behavior
- Chlorine treatment

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

- 632 upper bound
- 103 lower bound
- 333
- 319 lower bound
- 453 upper bound
- 158 lower bound
- 115
- 71
- 73 upper bound
- 66 lower bound
Developing an evaluation strategy

A. Start with a question
B. Verify the question hasn’t been answered
C. State a hypothesis
D. Design the evaluation
E. Determine whether the value of the answer is worth the cost of the evaluation

F. With key questions answered from impact evaluations, process evaluation can give your overall impact

G. A few high quality impact studies are worth more than many poor quality ones

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

<table>
<thead>
<tr>
<th>Component</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Needs Assessment</td>
<td>A. What is the problem?</td>
</tr>
<tr>
<td>Program Theory Assessment</td>
<td>B. How, in theory, does the program fix the problem?</td>
</tr>
<tr>
<td>Process Evaluation</td>
<td>C. Does the program work as planned?</td>
</tr>
<tr>
<td>Impact Evaluation</td>
<td>D. Were its goals achieved? The magnitude?</td>
</tr>
<tr>
<td>Cost Effectiveness</td>
<td>E. Given magnitude and cost, how does it compare to alternatives?</td>
</tr>
</tbody>
</table>