Theory of Change and Measurement

Ilf Bencheikh
Deputy Executive Director
J-PAL Europe
Lecture Overview

- Theory of change
- Sources of measurement
- Measurement concepts
Lecture Overview

- Theory of change
- Sources of measurement
- Measurement concepts
Components of Program Evaluation

- Needs Assessment
- Theory of Change
- Process Evaluation
- Impact Evaluation
- Cost Effectiveness Analysis
Definition

• The **Theory of Change** defines a set of connected building blocks, generally called inputs, outputs, and outcomes.

• It is depicted on a map that is a graphic representation of the change process.
Theory of Change

Needs Assessment

Intervention/Input

Output 1

Output 2

Intermediate outcome 1

Intermediate outcome 2

Intermediate outcome 3

Outcome 1

Outcome 2
Teachers work harder

Parents are informed about their kids learning

Parents get more involved in education at home
Parents get more involved in the School Committee
Teachers work harder

Better learning levels

Kids don’t learn and parents are not aware of it

Testing materials, field workers

THEORY OF CHANGE AND MEASUREMENT
No final impact: what can we conclude?

- Testing materials, field workers
- Parents are informed about their kids learning
- Parents get more involved in education at home
- Parents get more involved in the School Committee
- Teachers work harder
- Better learning levels
- Data

=> e.g: no impact
Scenario 1

Inputs

Testing materials, field workers

Outputs

Parents are informed about their kids learning

Intermediate Outcomes

Parents get more involved in education at home
Parents get more involved in the School Committee

Outcomes

Teachers work harder
Better learning levels
Scenario 2

- **Inputs**: Testing materials, field workers
- **Outputs**: Parents are informed about their kids learning
  - Intermediate Outcomes:
    - Parents get more involved in education at home
    - Parents get more involved in the School Committee
    - Teachers work harder
  - Outcomes: Better learning levels
Scenario 3

**Inputs**
- Testing materials, field workers

**Outputs**
- Parents are informed about their kids learning
  - Parents get more involved in education at home
  - Parents get more involved in the School Committee
  - Teachers work harder

**Intermediate Outcomes**
- Better learning levels
Scenario 4

- **Inputs**: Testing materials, field workers

- **Outputs**: Parents are informed about their kids learning (OK), Parents get more involved in education at home (OK), Parents get more involved in the School Committee (X), Teachers work harder (X)

- **Intermediate Outcomes**: Better learning levels (X)

- **Outputs**: Better learning levels (X)
Huge impact

**Inputs**
- Testing materials, field workers

**Outputs**
- Parents are informed about their kids learning
  - Parents get more involved in education at home
  - Parents get more involved in the School Committee
  - Teachers work harder

**Intermediate Outcomes**
- Better learning levels

**Outcomes**
- Data

=> e.g: huge impact

THEORY OF CHANGE AND MEASUREMENT
Scenario 1

**Inputs**
- Testing materials, field workers

**Outputs**
- Parents are informed about their kids learning

**Intermediate Outcomes**
- Parents get more involved in education at home
- Parents get more involved in the School Committee
- Teachers work harder

**Outcomes**
- Better learning levels
Scenario 2

**Inputs**
- Testing materials, field workers

**Outputs**
- Parents are informed about their kids learning
  - Parents get more involved in education at home
  - Parents get more involved in the School Committee
  - Teachers work harder

**Intermediate Outcomes**
- OK

**Outcomes**
- Better learning levels
Scenario 3

Inputs
- Testing materials, field workers
  - OK

Outputs
- Parents are informed about their kids learning
  - OK
- Parents get more involved in education at home
  - OK
- Parents get more involved in the School Committee
  - OK
- Teachers work harder
  - OK

Intermediate Outcomes
- Better learning levels
  - OK

Outcomes
- OK
You need data at each stage

Inputs
- Testing materials, field workers
  - Data

Outputs
- Parents are informed about their kids learning
  - Data
  - Intermediate Outcomes
    - Parents get more involved in education at home
      - Data
    - Parents get more involved in the School Committee
      - Data
      - Intermediate Outcomes
        - Teachers work harder
          - Data
    - Intermediate Outcomes
      - Data

Intermediate Outcomes
- Parents get more involved in education at home
  - Data
- Parents get more involved in the School Committee
  - Data
- Teachers work harder
  - Data

Outcomes
- Better learning levels
  - Data
Why spend time on the theory of change?

1. Helps design the intervention
   - Often done backwards
   - Is each arrow really credible?

2. Helps design the evaluation by:
   - Generating research questions
   - Deciding which data to collect

   By measuring the right intermediate variables, we can get “into the black box”
   - Allows to understand the “why”, thus giving richer policy lessons
   - Gives more generalizable knowledge
Lecture Overview

- Theory of change
- Sources of measurement
- Measurement concepts
Where can we get data?

- From existing sources (Secondary data)
  - Publicly available
  - Administrative data
  - Other secondary data
- Collected by researchers (Primary data)
  - Surveys
  - Non-survey methods
## Types and Sources of Data

<table>
<thead>
<tr>
<th>Information reported by a person</th>
<th>Information about a person/household</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognition, anthropometrics</td>
<td>Cognition, anthropometrics</td>
</tr>
<tr>
<td>Demographics</td>
<td>Demographics</td>
</tr>
<tr>
<td>Behavior, beliefs</td>
<td>Behavior, beliefs</td>
</tr>
<tr>
<td>Patience, risk aversion,</td>
<td>Patience, risk aversion, psychometrics</td>
</tr>
<tr>
<td>Knowledge</td>
<td>Knowledge</td>
</tr>
<tr>
<td>Income, expenditure</td>
<td>Income, expenditure</td>
</tr>
<tr>
<td>Bank transactions</td>
<td>Bank transactions</td>
</tr>
<tr>
<td>Phone data</td>
<td>Phone data</td>
</tr>
<tr>
<td>Sales records</td>
<td>Sales records</td>
</tr>
<tr>
<td>School/university records,</td>
<td>School/university records, criminal record</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Automatically generated</th>
<th>NOT about a person/household</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farming inputs and outputs</td>
<td>Farmsing inputs and outputs</td>
</tr>
<tr>
<td>Quality of medical care</td>
<td>Quality of medical care</td>
</tr>
<tr>
<td>Business income taxes</td>
<td>Business income taxes</td>
</tr>
<tr>
<td>Prices</td>
<td>Prices</td>
</tr>
<tr>
<td>Weather, air quality</td>
<td>Weather, air quality</td>
</tr>
<tr>
<td>Stock markets</td>
<td>Stock markets</td>
</tr>
<tr>
<td>VAT records</td>
<td>VAT records</td>
</tr>
</tbody>
</table>
Primary Data Collection

- Surveys
- Exams, tests, etc.
- Games
- Vignettes
- Direct observation
- Diaries/logs
- Focus groups
- Interviews
Primary Data: Modes

• Interviewer administered
  – Paper-based
  – Computer-assisted/ digital
  – Telephone-based

• Self-administered
  – Paper
  – Computer/digital
Administrative data

Information collected, used, and stored primarily for administrative (i.e., operational) purposes, rather than research purposes.

- Medical records
- Grade books
- Arrest records
- Bank account data
- Personnel records
- Log books
Why are administrative data useful?

The outcomes and metrics required for a study may already be tracked by a government or organization

- Available retrospectively
- Enable long-term follow-up
- May include near census of relevant population
- Reduce logistical burden and burden on subjects
- Often less expensive than surveys
- May reduce bias and error
Lecture Overview

• Theory of change

• Sources of measurement

• Measurement concepts
The main concept being investigated. A construct is often abstract.
(E.g., learning, teacher effort, parental involvement)

How you actually measure or “operationalize” your construct.
(e.g. test scores, teacher preparedness, hours dedicated to helping kids do homework)

What we use to measure our indicators.
Validity

In theory: how well does the indicator map to the construct we are trying to measure?
Is the indicator measured in a way that is consistent and precise?

Would we get the same data if we measured several times?
The Goals of Measurement

Reliability

Low

High

Validity

Low

High
Data collection can go wrong

Construct → Indicators → Data Collection → Data → Measurement error / bias
The Response Process

1. Comprehension of the question
2. Retrieval of Information
3. Judgement and Estimation
4. Reporting an Answer
Step 1: Comprehension

1. Comprehension of the question
2. Retrieval of Information
3. Judgement and Estimation
4. Reporting an Answer

1.1 Total monthly income, before taxes
____________________
____________________
____________________
____________________
Step 2: Retrieval

1. Comprehension of the question
2. Retrieval of Information
3. Judgement and Estimation
4. Reporting an Answer

Social Security benefits, Unemployment or Workers’ Compensation, Pensions...
Step 3: Estimation/Judgement

1. Comprehension of the question
2. Retrieval of Information
3. Judgement and Estimation
4. Reporting an Answer

Social = $200 per month
Workers’ Compensation = 0
Pension = $220 per month
What else??
Step 4: Response

1.1 Total monthly income, before taxes

____________________

____________________

____________________

____________________

$400
Conclusion on data quality

• Particularly difficult to ensure the quality of survey data
  => Lot of energy and resources dedicated to that

• Good to combine survey data with other kind of data (direct observation, tests)...

• But those too can be imprecise / biased

• Admin data potentially subject to similar imprecision / biases as primary data
Time to illustrate with a case study!
Thank you!