Measuring Impact
Outcomes & Indicators
01 March, 2013

Diva Dhar
CLEAR South Asia at J-PAL South Asia at IFMR
Presentation Overview

- **Indicators:** What makes a good measure?
- Theory of Change: What do you want to measure?
- Data Sources: How could you measure it?
- Data Collection: How should you measure it?
The main challenge in measurement: getting accuracy and precision

More accurate →

More precise
Accuracy

- How well does the indicator map to the outcome?
  - Standardized Tests, Income, empowerment

- Are you getting unbiased answers?
  - Social desirability bias (response bias)
  - Framing effect
  - Recall bias
  - Anchoring bias
  - Neutrality
  - Respondent Bias
Which is worse?

A. Poor accuracy
B. Poor precision
C. Equally bad
D. Depends
E. Don’t know/can’t say

A. 18%
B. 32%
C. 27%
D. 23%
E. 0%
Accuracy AND Precision

- Question wording
- Surveyor training/quality
- Data entry
- Translation
- Length, fatigue
Precision and Error

- Question wording
  - Definition of terms – E.g. ‘household’, ‘income’
  - Recall period / units of question

- Type of answer
  - Open/Closed
  - Choice of options for closed questions
    - Likert (i.e. Strongly disagree, disagree, neither agree nor disagree, . . .)
    - Rankings
    - Quantitative scales. Numbers or bins.
How frequently do you drive your car?

A. Never
B. 1-2
C. 3-5
D. 6-10
E. More than 10
F. Don’t know/can’t say
The Problem

- With the following questions…
Outcome: Annual consumption
Indicator: Food expenditure in last week
(The problem with this indicator is....)

A. Accuracy
B. Precision
C. Both
D. Neither
Outcome: Annual food consumption
Indicator: Food expenditure in last three months
(The problem with this indicator is....)

A. Accuracy
B. Precision
C. Both
D. Neither
Question: After the last time you used the toilet, did you wash your hands?
(The problem with this indicator is....)

A. Accuracy
B. Precision
C. Both
D. Neither
Presentation Overview

- **Indicators:** What makes a good measure?
- **Theory of Change:** What do you want to measure?
- **Data Collection:** How could you measure it?
- **Survey Design:** How can I measure it well?
Why do you need data? Follow Theory of Change

- Our theory and hypothesis helps us define the set of outcomes
- Need to find indicators that map the outcomes well
- Characteristics: Who are the people the program works with, and what is their environment?
  - Sub-groups, covariates, predictors of compliance
- Channels: How does the program work, or fail to work?
- Outcomes: What is the purpose of the program? Did it achieve that purpose?
Theory of Change: What do we want to achieve? What does that require?

- What do we want to achieve?
  - Purpose of the program, not the program itself (e.g., “educate children”, or “hand out textbooks”)

- What does that require?
  - What are the causal links that must exist for this program to achieve its purpose?
  - What assumptions are we making in defining these links, and which do we want to measure?
Theory of Change: Measurement for Spandana

**Situation/Context Analysis:** Low income, high debt, gender inequality

Access to business networks

Open Spandana branches targeting women

After 6 months, 80% of planned branches opened and operational.

Access to a branch

Eligible for a loan

Apply for a loan

Women take out loans

After 9 months, 20% of eligible women have taken out a loan at a Spandana branch.

Access to business networks

Know which businesses are profitable

Choose to invest in new or existing businesses

After 12 months, 25% of the women who have taken out a loan have invested in a business.

Women have entrepreneurial skills

Businesses are profitable

After 18 months, 50% of the businesses started by women w/ Spandana loans are profitable.

No external shocks

Higher incomes

After 18 months, 40% of the women who have started businesses with Spandana loans have incomes at least 10% higher.

After 6 months, 80% of planned branches opened and operational.

After 9 months, 20% of eligible women have taken out a loan at a Spandana branch.

After 12 months, 25% of the women who have taken out a loan have invested in a business.

After 18 months, 50% of the businesses started by women w/ Spandana loans are profitable.

After 18 months, 40% of the women who have started businesses with Spandana loans have incomes at least 10% higher.
# Logical Framework

<table>
<thead>
<tr>
<th>Impact (Goal/Overall objective)</th>
<th>Outcome (Project Objective)</th>
<th>Outputs</th>
<th>Inputs (Activities)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher income</td>
<td>Households start new businesses; expand existing ones</td>
<td>Increased MFI borrowing</td>
<td>MFI branches are opened</td>
</tr>
<tr>
<td>Spending</td>
<td>Purchase of durable goods</td>
<td>Number of microloans</td>
<td>Branches are operating; providing services</td>
</tr>
<tr>
<td>Household survey</td>
<td>Household survey</td>
<td>Household survey, Administrative data from MFIs</td>
<td>Branch visits/surveys</td>
</tr>
<tr>
<td>Business well-run, are profitable, no external shocks</td>
<td>No problems of self-control, time-inconsistency. HHs can identify good investment opportunities</td>
<td>Women have access to branches. No borrowing from informal sources</td>
<td>Sufficient technical and human resources, funding</td>
</tr>
</tbody>
</table>
Presentation Overview

1. Indicators: What makes a good measure?

2. Theory of Change: What do you want to measure?

3. Data Collection: How could you measure it?

4. Survey Design: How can I measure it well?
When to Collect Data

- [Baseline]: Before you start
- During the intervention
- Endline: After you’re done
- [Scale-up, intervention]
Methods of data collection

- Administrative data
- Surveys - household/individual
- Logs/diaries
- Qualitative – eg. focus groups, RRA
- Games and choice problems
- Observation
- Health/Education tests and measures
Where can we get data?

<table>
<thead>
<tr>
<th>Administrative data</th>
<th>The good . .</th>
<th>. . and the bad</th>
</tr>
</thead>
</table>
| • Collected by a government or similar body as part of operations | • May already be collected and thus free  
  • Can be extremely accurate (e.g. electricity bills) | • May not exist or not answer the question you want  
  • May itself change behavior (e.g. taxes) |
| Other secondary data                                      |              |                                                                                |
| • Collected for research or other purposes not admin      | • May already be collected and thus free  
  • Can inform the larger context of a project           | • May not exist or not answer the question you want  
  • Dubious quality                                      |
| Primary data                                              |              |                                                                                |
| • Collected by researchers for study                      | • Address the exact question of interest  
  • Cover channels and assumptions                         | • Very costly and time consuming  
  • May be biased answers                                  |
Primary Data Collection Considerations

- Quality
  - Reliability and validity of the data

- Costs / Logistics
  - Surveyor recruitment and training
  - Field work and transport, interview time
  - Electronic vs. paper
  - Data entry, reconciliation, cleaning, etc.

- Ethics
  - Human subjects, data security
Common survey modules can be adapted for a particular project

- Demographics
- Economic
  - Income, consumption, expenditure, time use
  - Yields, production, etc.
- Beliefs
  - Expectations or assumptions
  - Bargaining power, patience, risk
- Anthropometric
- Cognitive, Learning
Presentation Overview

1. Indicators: What makes a good measure?
2. Theory of Change: What do you want to measure?
3. Data Collection: How should you measure it?
4. Survey Design: How can I measure it well?
Outcome: Gender Bias
Question: How effective are women leaders? (ineffective, somewhat effective, effective, very...)

A. Accuracy
B. Precision
C. Both
D. Neither

A. 22%
B. 35%
C. 26%
D. 17%
Perceptions and Attitudes

- Ask directly
  - “How effective is your leader?” (ineffective, somewhat effective, effective, very...)

- Indirect approaches often have better accuracy
  - Listen to a Vignette (Male v. Female)
  - Revealed preference – voting behavior
  - Implicit Association tests
Implicit Association Test
Find a match on Left or Right?

Women or Sport

89%

Men or Music

11%

Cricket
Find a match on Left or Right?

Women or Music

Men or Sport

100% Cricket

0% Cricket
Implicit Association Test

- People simplify the world for efficiency
  - Use thumb rules to draw connections
  - May not even be aware themselves

- For some important outcomes, may be worth trying to measure these indirectly
  - Implicit association one technique
  - Actually based on response time, not accuracy
Asking sensitive questions: List randomization

- Some questions are sensitive and people are hesitant to answer truthfully

- List randomization is a way to get the answer on average *without* knowing confidential information on any one person.
Asking sensitive questions: List randomization

- Randomly ask part of the sample the question with / without a sensitive option
- Response only a count, not the specific options

How many of these statements are true for you?

- This morning I took a shower.
- My nearest bank branch office is within walking distance.
- I have tea every day.

How many of these statements are true for you?

- This morning I took a shower.
- My nearest bank branch office is within walking distance.
- I have tea every day.
- I use my loan for non-business expenses.
Asking sensitive questions: List randomization

2.8 – 2.1 = 0.7 full TRUE difference (on average)

70% used their loan for non-business purposes.

Average number of true statements:

2.1

2.8

Average number of true statements:
List randomization shows big differences in some real cases.
Reliability of Data Collection

- The process of collecting “good” data requires a lot of efforts and thought
- Need to make sure that the data collected is precise and accurate.
  - avoid false or misleading conclusions
- The survey process:
  - Design of questionnaire → Survey printed on paper/electronic → filled in by enumerator interviewing the respondent → data entry → electronic dataset
- Where can this go wrong?
Reliability of Survey Data

- Start with a pilot
- Paper vs electronic survey
- Surveyors and supervision
- Following up the respondents
- Problems with respondents
- Paper or electronic version?
- Neutrality
Things to Take-away

- Theory of change guides measurement
- Data collection all about trade-offs
  - Quality and cost
  - Validity (accuracy) and reliability (precision)
- Creative techniques can sometimes help
  - Think about what outcomes are most important