Instrument Design 101

Diva Dhar
Policy and Training Manager
J-PAL at IFMR
Outline of lecture

- Methods of data collection
- Deciding what to ask
- Deciding who to ask
- Writing and wording questions
- Structuring and formatting
- Questionnaire design resources
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
Conversation with concerned individuals

Field visits

Community interviews

Reviews of official records (management information and system and administration data)

Participant observations

Key informant interviews

Focus group interviews

Direct observations

Survey

Panel surveys

One time surveys

Censuses

Field experiments

Informal/ less structured

Formal/ more structured
Deciding what to ask: Purpose of the Questionnaire

- Each question should relate to the study and theory of change
  - Covariates
  - Measurement outcomes

- List all indicators for what you intend to measure
  - Use participatory approach to develop indicators (existing instruments, experts, beneficiaries, stakeholders)
  - Assess based on feasibility, time, cost and importance
Deciding what to ask: Modules

- “Modules” = components or sub-sections of instrument
- Each module captures a different theme
- Logical order, from
  - general -> specific,
  - factual -> attitudinal/behavioral,
  - least sensitive -> most sensitive
- Watch out for priming and question order
Choosing Modules: Units of Observation

Often this is simple: For example, sex and age are clearly attributes of individuals. Roofing material is attribute of the dwelling.

Not always obvious: To collect information on credit, one could ask household’s

1. All current outstanding loans.
2. All loans taken and repaid in the last one year.
3. All “borrowing events” (all the times a household tried to borrow, whether successfully or not).

Each of the three units of observation has its analytic benefits…and biases
Deciding who to ask: Target respondent

- “Target respondent”: should be most informed person for each module. Respondents for each module can vary.

- For example: to measure use of TLM, should we survey the headmaster? Teachers? SMC? Parents? Students?

- Choice of modules decides target respondent, and target respondent shapes module design.
Deciding how to ask: Writing and Adapting Questions

A. Type of question
B. Wording of the question
C. Time-frame/Recall Periods

A. Estimation Technique
B. Unit of Measurement
C. Detailed coding of answer choices
A) Types of Questions

- Close-ended versus Open-ended
- Single vs. Multiple Response
- Ranking
- Rating
A) Types of Questions

- Open-Ended versus Close-Ended:
  - Why don’t you send your child to school?

- Gives precise judgments;
  But...if you are concerned, as most surveys are, in summarizing replies to produce picture of population, better approach is probably a closed ended question:

  OR

- Do you not send your child to school because:
  - Nearest school is too far away
  - Teacher does not come
  - Need child to work in the field
A) Types of Questions

- Single vs. Multiple Response
  Do you want respondent to give single or multiple answers?

- Question formats: Likert ("Rating") Scales
  Example: Boys should be allowed to get more opportunities and resources for education than girls (strongly agree, agree, neutral, disagree, and strongly disagree)

AVOID asking descriptive terms for facts: ask for numerical estimate.

- Inaccurate: How often do you deposit savings in your account?
  Never, hardly ever, sometimes, constantly

- Accurate: In the last 1 month, how many times did you deposit savings?
  Number of times______________.
A) Types of Questions

- Ranked Responses

Sometimes useful for respondent to rank a set of options by numbering

**Example**: Which Teaching-Learning Materials do you use?

**Options**: Workbooks, flipcharts, textbooks, maps, flash cards, games etc.

Which teaching-learning materials do you use most often?

*(Rank the top 3)*

**Trouble spots:**

- Assumes respondent feels differently about each item
- Assumes respondent knows enough about each item to make a comparative judgment
B) Clear definitions and Precise wording

How do you define a household?

- Is it “A 'household' is usually a group of persons who normally live together and take their meals from a common kitchen/common cooking unless the exigencies of work prevent any of them from doing so” as in the Census of India

- Is it “A 'household' consists of two types of people: 1) people who usually live in the household and 2) people who stayed in the household the night before the day of the survey.” as in the DLHS
B) Clear definitions and Precise wording

- How do you define income?
  - Any activity with compensation?
  - What if compensation is in the form of rice/wheat/maize?
  - Rent?
  - Pension or retirement income?
  - Welfare payments?
B) Clear definitions and Precise wording

**Question:** “What proportion of children do not know the basics?”

**Questions:** “Do officers monitor the schools regularly?’

Wording of the questions should be precise, not ambiguous
C) Recall Periods

Trade off potential recall (memory) bias from long period reporting against potential variance from short periods

Example, consumption:

- Frequently purchased items (e.g. food), recall period between one week and one month
- When practical: diaries for a week or two
- Durable goods: annual recall basis

→ Reference periods can be different across questions, and sections

→ Different recall periods can present a problem when combining or comparing data across modules
D) Estimation

Estimation is important, but can be tricky

- Provide tools and benchmarks
- Break down the calculations
- Cross-check with multiple sources
E) Detailed Response Coding

- Designing good response codes requires strong knowledge of the phenomena being studied.
  
  - Background research and possible qualitative research are handy here. See similar questionnaires.
  
  - Ideally responses to almost all questions should be given code numbers, and only code numbers, as opposed to words or phrases, should be recorded on the questionnaire and entered.
E) Detailed Response Coding

- Specifying Wording – Pre-coding

Pre-coding requires that codes be clear, simple, and:

1. **mutually exclusive**;
2. that they **exhaust all likely answers**,
3. that **respondents will not all provide the same response**,
4. and that **none of the codes apply to only a handful of respondents**

To ensure codes encompass all possible answers, add “other (specify______)” code to questions where precoding each possible response is impossible/ inconvenient.
## Structuring and Formatting

- **Skips**
- **Surveyor Instructions**

<table>
<thead>
<tr>
<th>Question</th>
<th>Options</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2.13 Has this person ever attended school (formal or informal)?</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>2</td>
</tr>
<tr>
<td>A2.14 What is the highest class of school this person has completed?</td>
<td>Did not complete Class 1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Class 1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Class 2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Class 3</td>
<td>3</td>
</tr>
<tr>
<td>A2.15 How many days of school did the person attend in the last 14 days</td>
<td>Record Number</td>
<td></td>
</tr>
<tr>
<td>A2.15a Did the person miss any days of school due to illness</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Will not answer</td>
<td>-888</td>
</tr>
<tr>
<td></td>
<td>Does not know</td>
<td>-999</td>
</tr>
<tr>
<td>A2.16 INTERVIEWER CHECKPOINT: Is this person 13 years old or older?</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>2</td>
</tr>
</tbody>
</table>
## Structuring and Formatting

### Consistent styles

<table>
<thead>
<tr>
<th>Question</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1.14 Is he/she studying this year or going to angawadi/balwadi?</td>
<td>YES 1, NO 2, DO NOT KNOW -999</td>
</tr>
<tr>
<td>A1.15 What is his/her marital status?</td>
<td>MARRIED, LIVING WITH SPOUSE 1, MARRIED, NOT LIVING WITH SPOUSE YET 2, SEPARATED/DIVORCED 3, SPOUSE/PARTNER DIED 4, N/A 5, NEVER MARRIED DONE N/A 6</td>
</tr>
<tr>
<td>A1.16 How many spouses and/or partners does this person have now?</td>
<td></td>
</tr>
<tr>
<td>A1.17 Do any of these spouses and/or partners live in this house?</td>
<td>YES 1, NO 2</td>
</tr>
<tr>
<td>A1.18 What is the name(s) of his/her spouse(s) and/or partner(s)?</td>
<td>INTERVIEWER: WRITE IN CODE NUMBER(S) FROM BOARD</td>
</tr>
<tr>
<td>A1.19 In what year, or at what age, did he/she first get married or do a 'nata'?</td>
<td>YEAR, AGE, DO NOT KNOW -999</td>
</tr>
<tr>
<td>A1.20 Is this person mentally retarded or disturbed?</td>
<td>YES 1, NO 2</td>
</tr>
<tr>
<td>A1.21 Is this person handicapped?</td>
<td>YES 1</td>
</tr>
</tbody>
</table>
Questionnaire Design Resources

- Please don’t forget to grab the hand-outs!
- Guidelines on design and piloting
- List of helpful books, papers
- Literature reviews and survey questions
THANK YOU!

Questions?