Measurement

Outcomes, Indicators, Data

Harini Kannan
Post Doctoral Fellow and Senior Research Manager
J-PAL South Asia at IFMR

July 2017

Lecture Overview

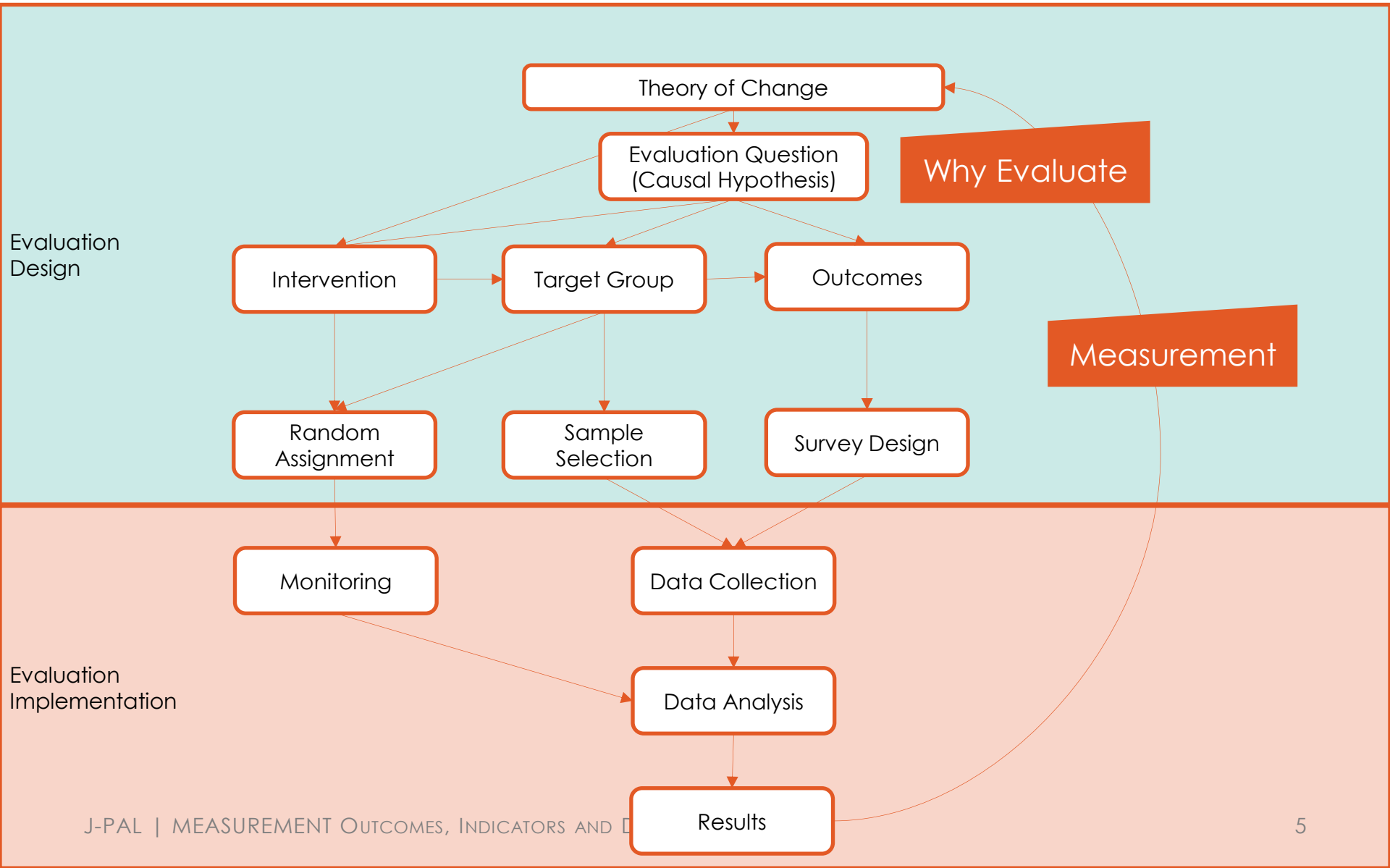
1. **What to Measure**
   - Theory of Change and Indicators
   - Measurement Concepts
2. **Challenges in Measurement**
   - Response Process
   - Measurement Error
3. **How to Measure It (Well)**
   - Sources of Measurement
4. **Concluding Thoughts**

What to Measure

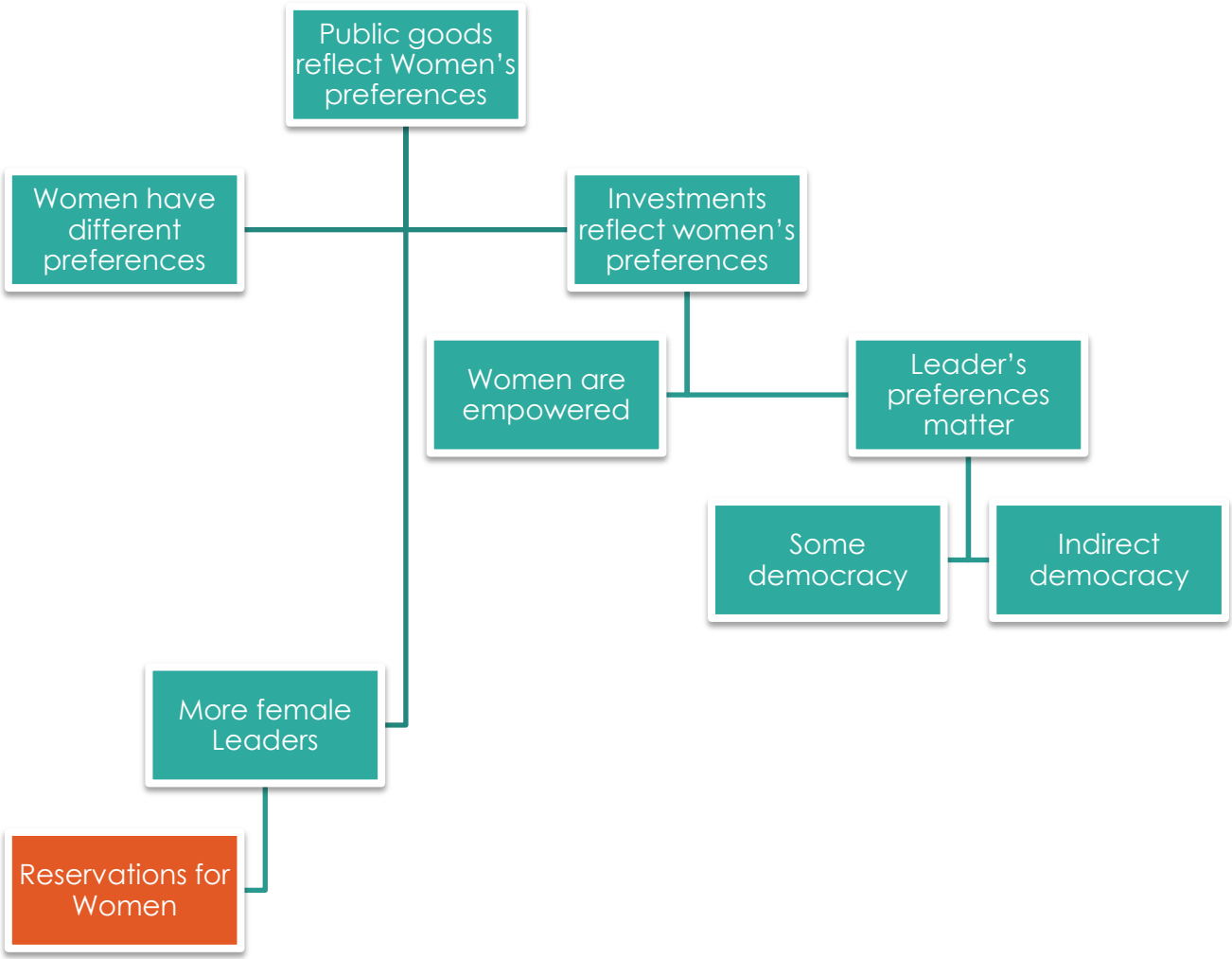
Theory of Change and Indicators

Measurement Concepts

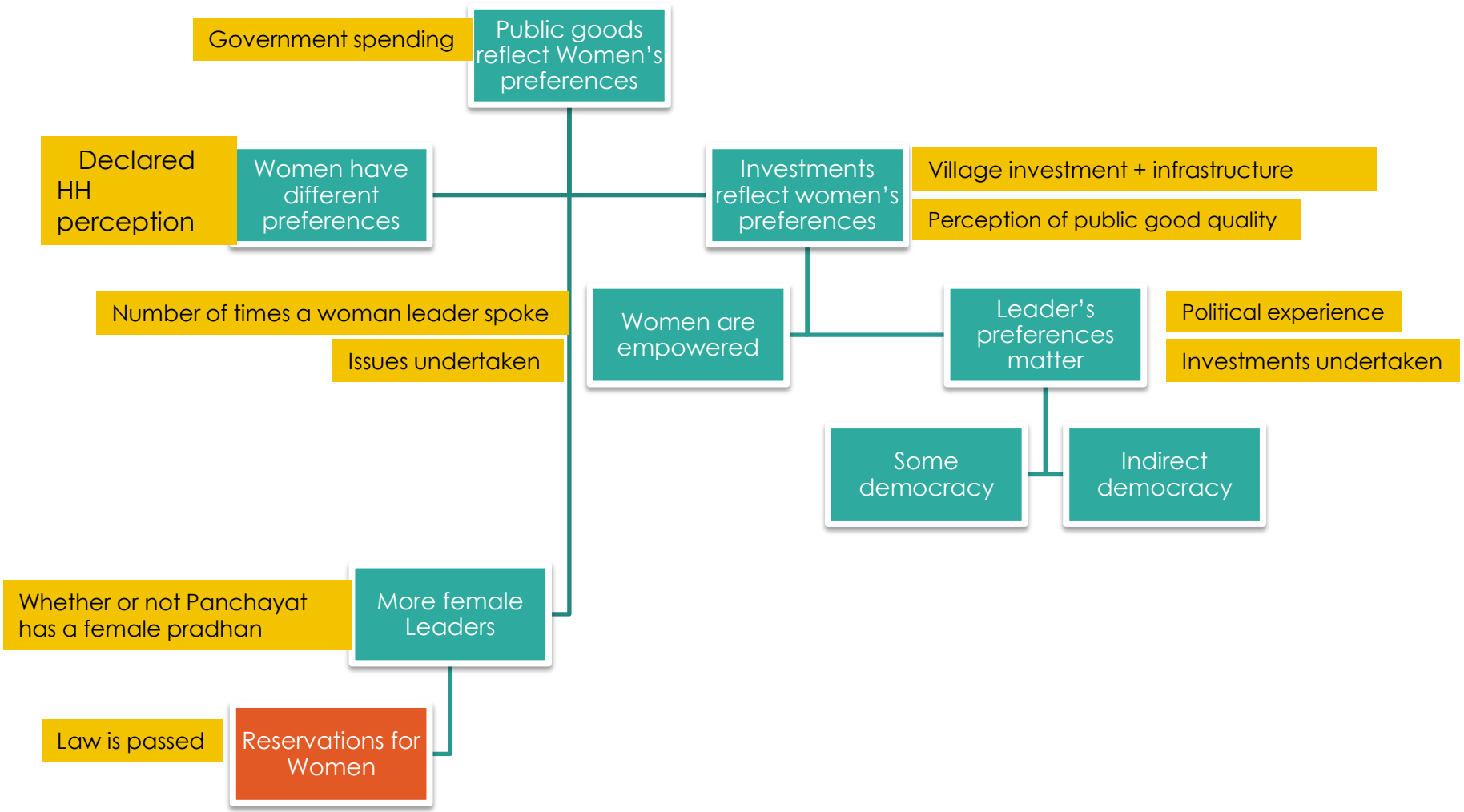
Randomized Evaluation Process



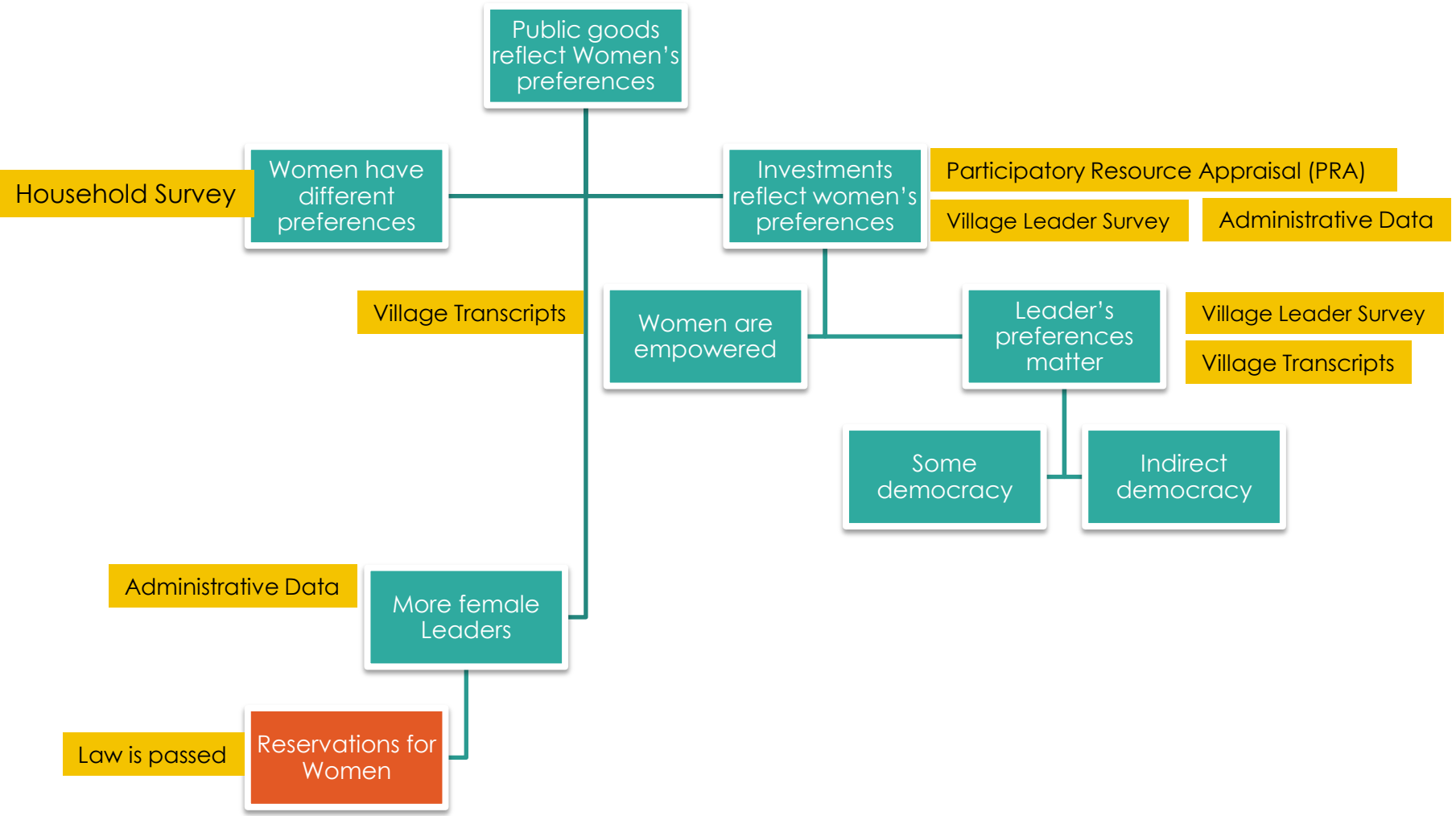
Theory of Change



Theory of Change: What to measure?

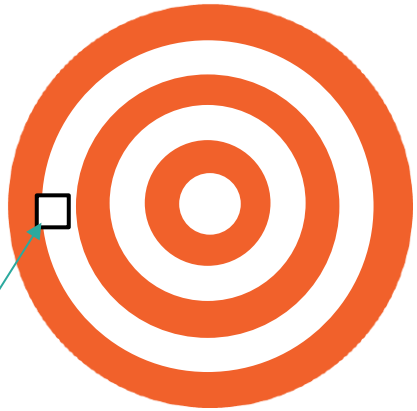


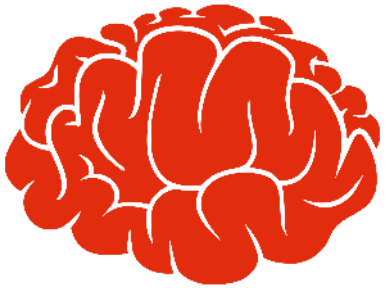
Theory of Change: How to measure?



Concept of measurement

Concept of measurement

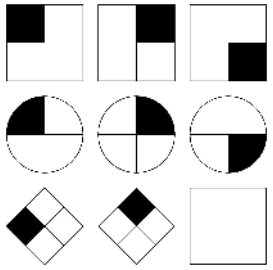




Data (Test Result)

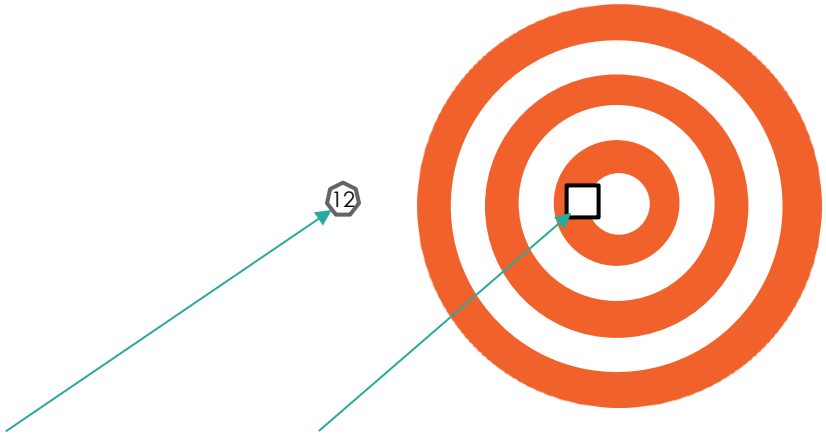
Indicator (IQ Test)

Construct (Intelligence)

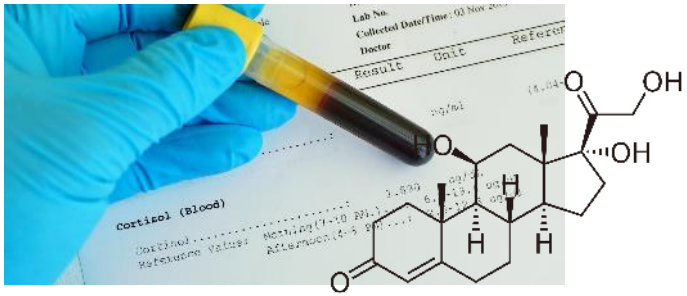


https://commons.wikimedia.org/wiki/File:Red_Silhouette_-_Brain.svg

Concept of measurement



Data (Test Result) Indicator (Cortisol level) Construct (Stress)



https://pixabay.com/en/despair-stress-alone-being-alone-862349/

Kilograms of rice per hectare:

A. A construct
B. An indicator
C. A response
D. Data
E. Don't know

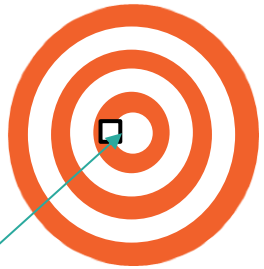
Empowerment is:

A. A construct
B. An indicator
C. A response
D. Data
E. Don't know

“Blood Pressure = 110/71 mm Hg” is:

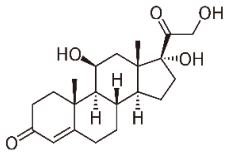
A. A construct
B. An indicator
C. A response
D. Data
E. Don’t know

The goals of measurement

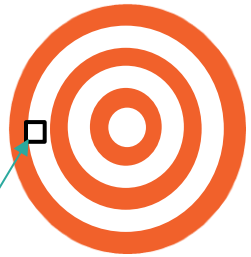


Data
(Test Result)

Indicator
(Cortisol level)

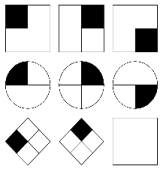


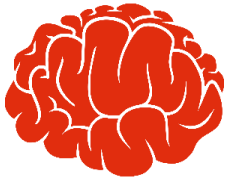
Construct
(Stress)



Data
(Test Result)

Indicator
(IQ Test)

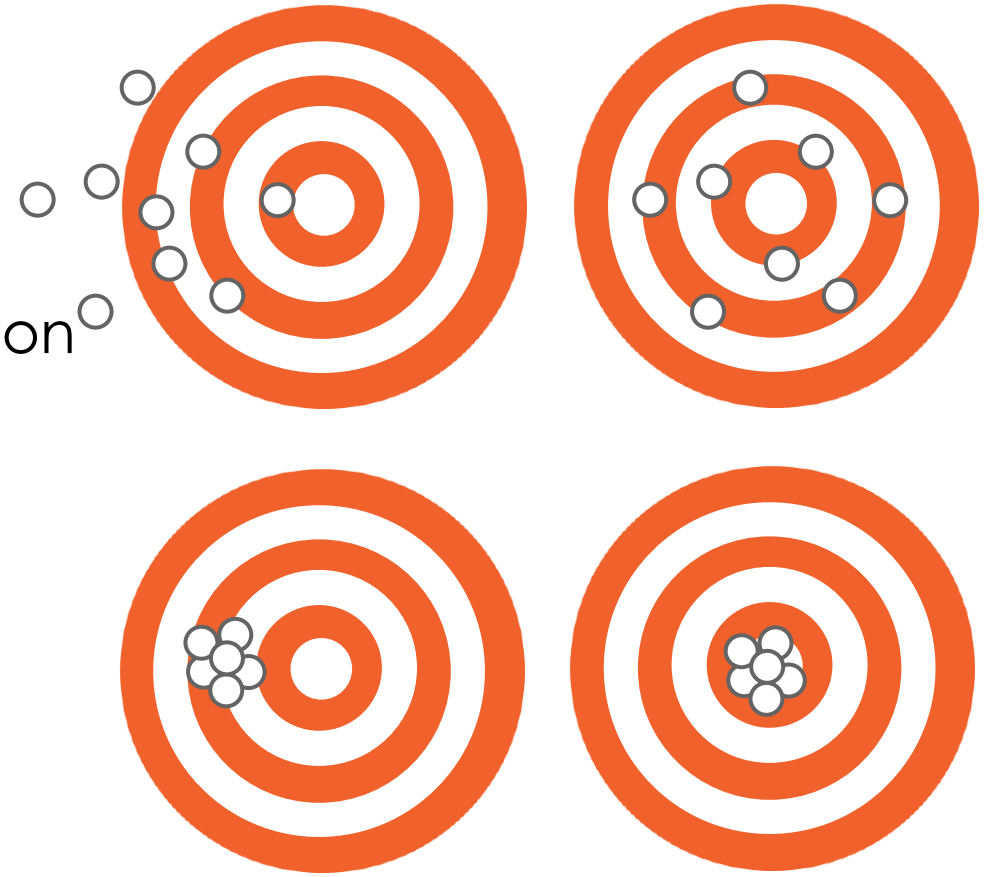




Construct
(Intelligence)

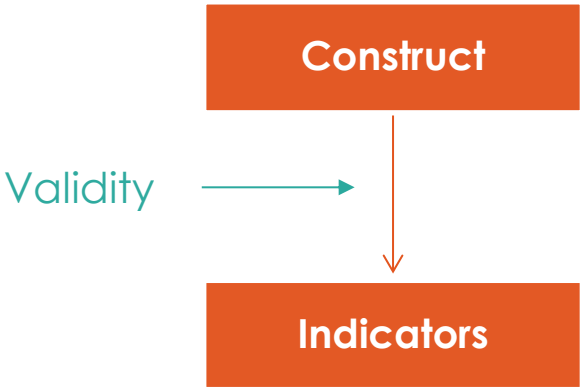
The goals of measurement

- Validity
- Unbiasedness
- Accuracy/Precision
- Reliability



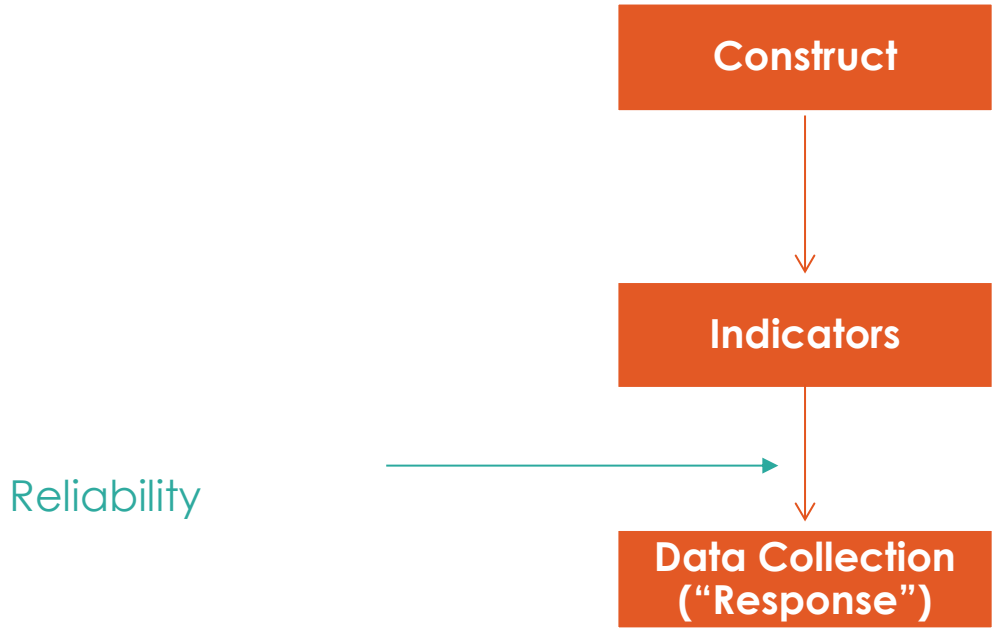
Validity

- In theory:
  - How well does the indicator map to the outcome? (e.g. IQ tests → intelligence)



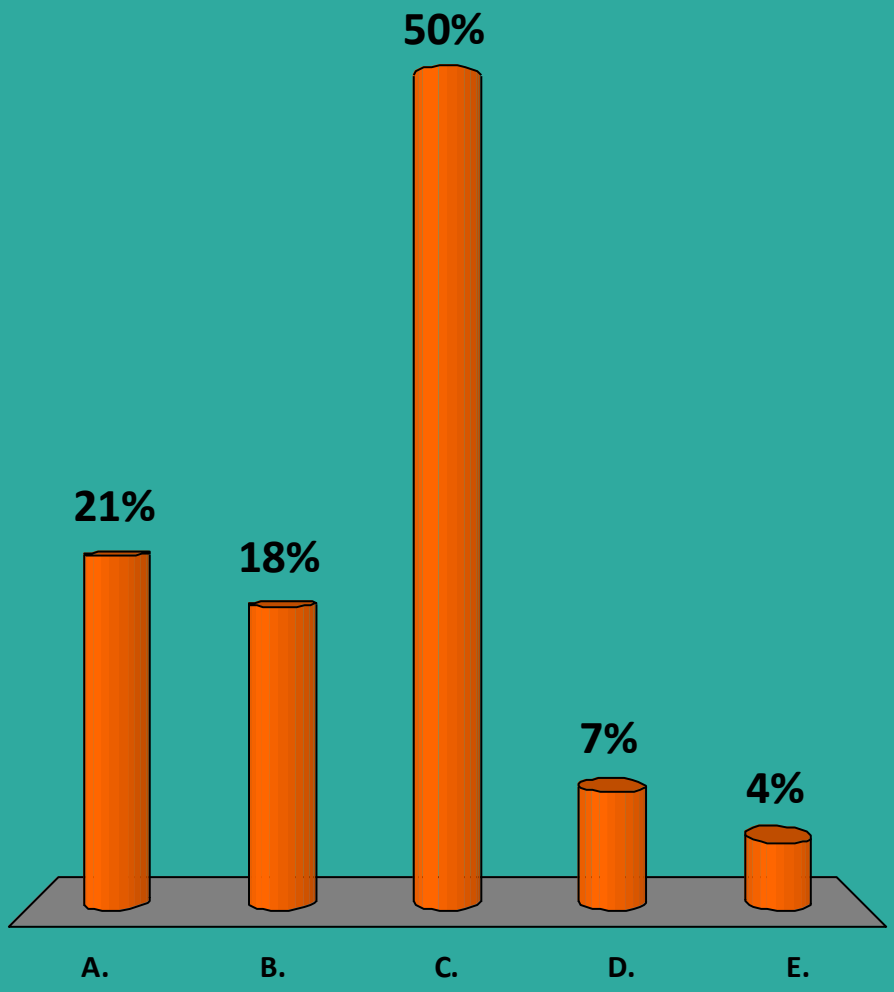
Reliability

- In theory:
  - The measure is consistent and precise vs. "noisy"



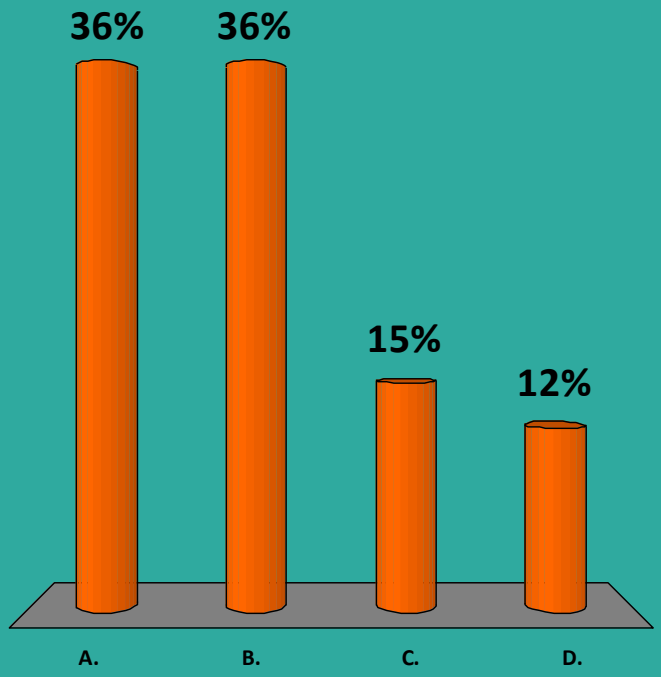
Which is worse?

A. Poor Validity
B. Poor Reliability
C. Equally bad
D. Depends
E. Don't know/can't say



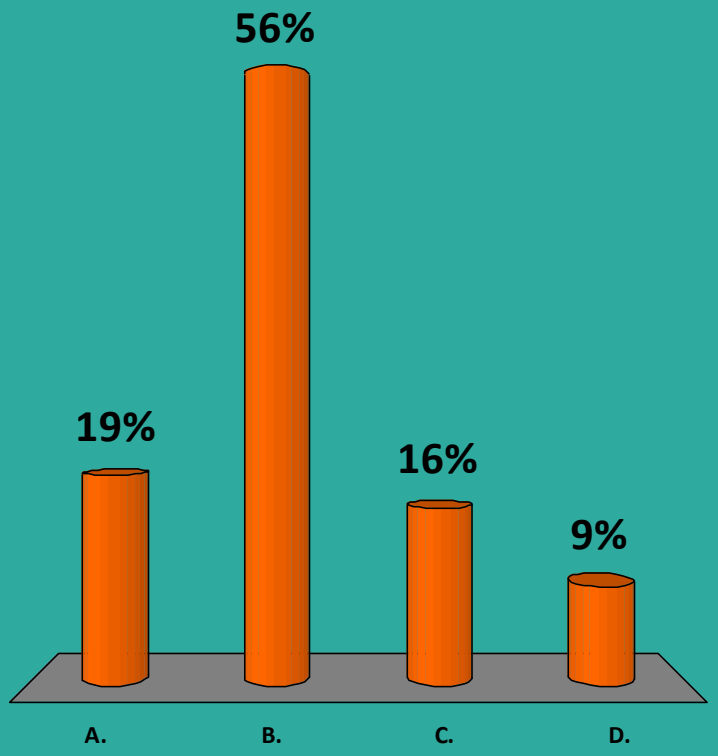
Outcome: annual consumption
Indicator: food expenditure in last week

A. Validity
B. Reliability
C. Both
D. Neither



Outcome: annual consumption
Indicator: food expenditure in last three months

A. Validity
B. Reliability
C. Both
D. Neither

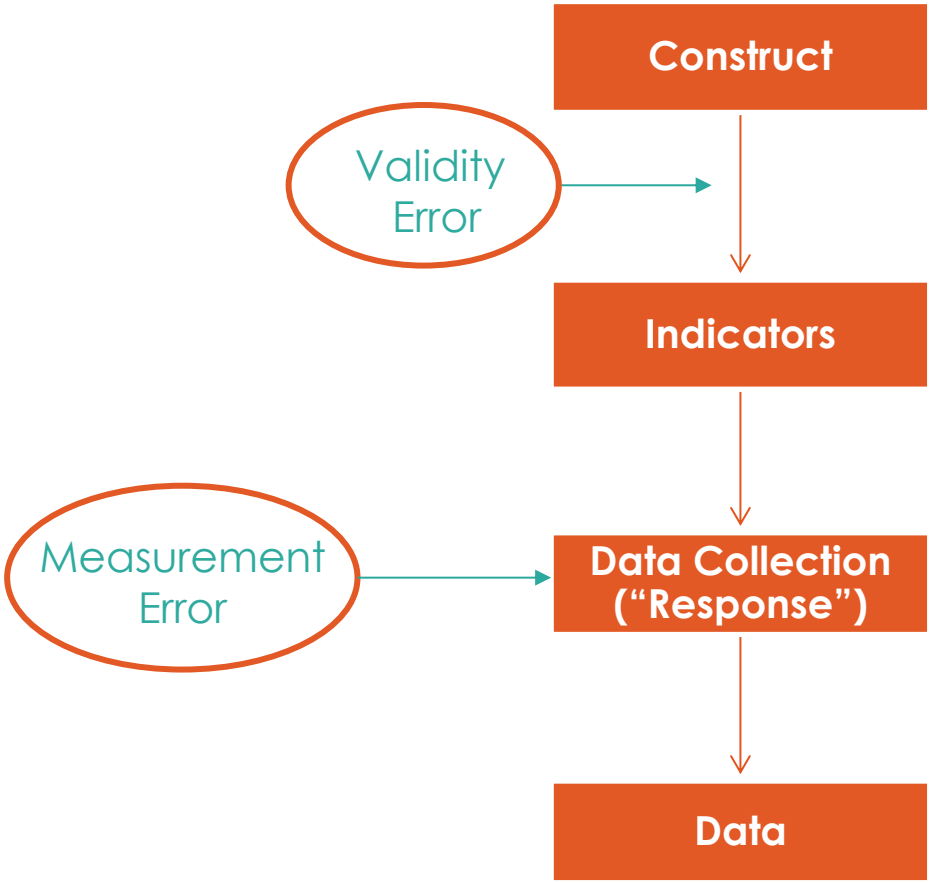


Challenges in Measurement

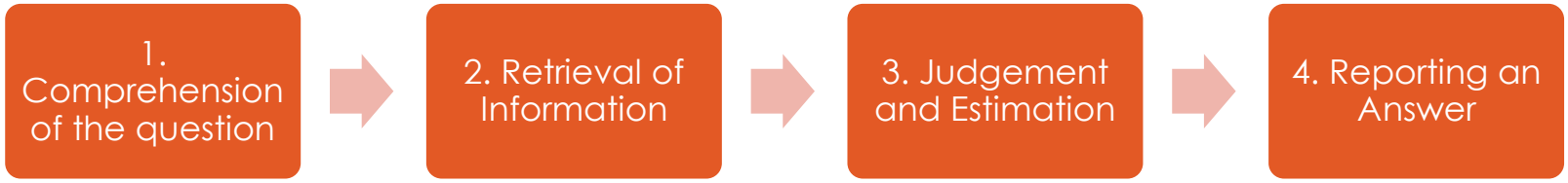
Response Process

Measurement Error

Error in Measurement



4-step Response Process



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

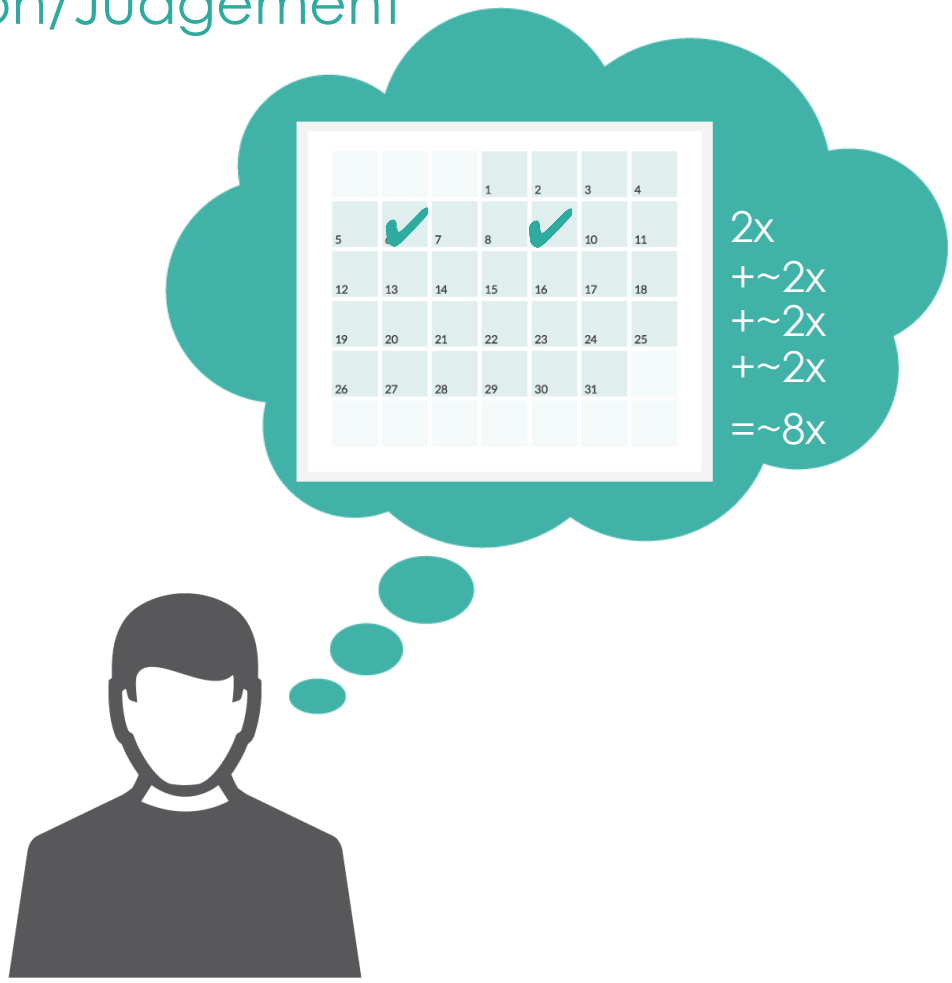
Step 1: Comprehension

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

Step 2: Retrieval

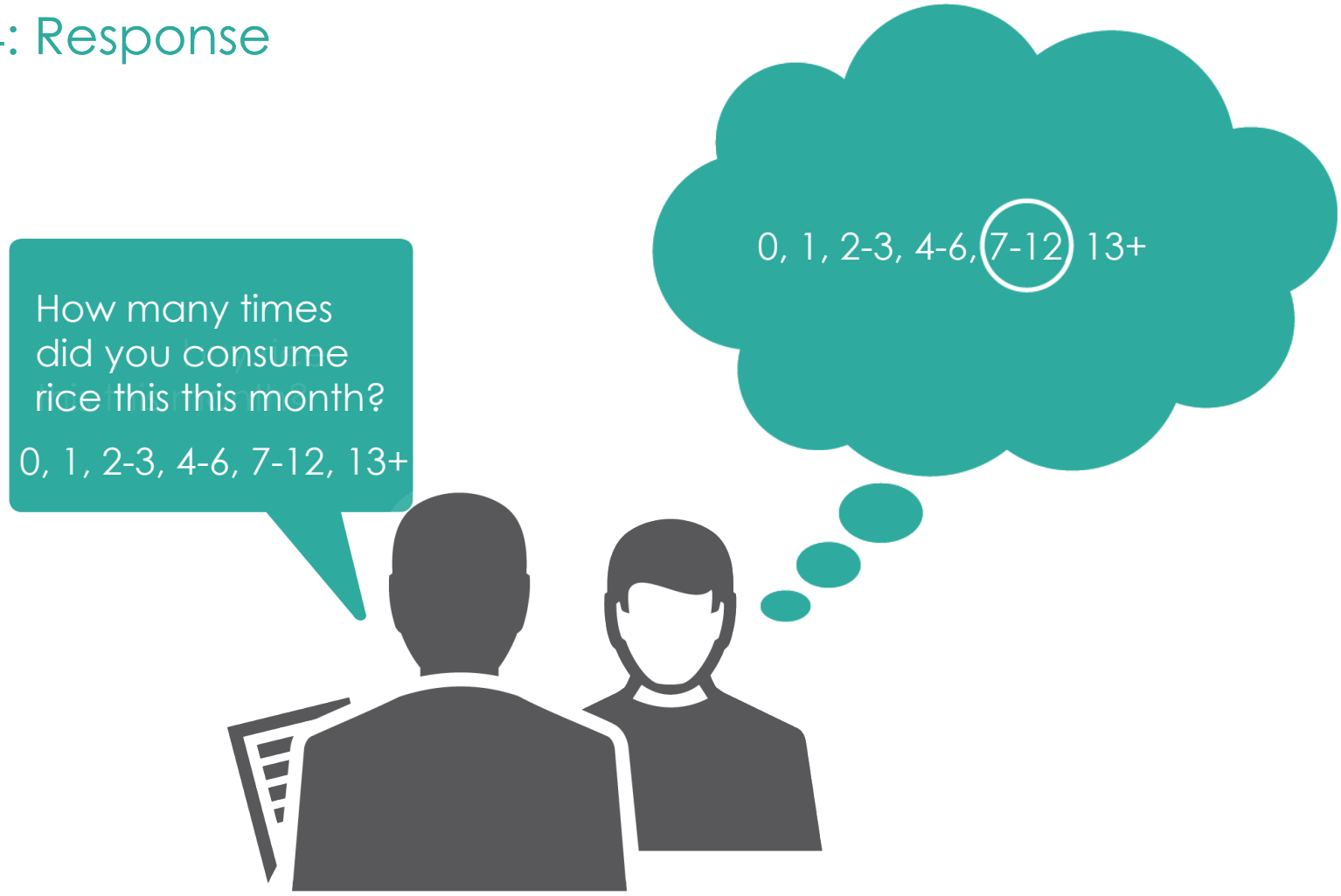
1. Comprehension of the question → 2. Retrieval of Information → **3. Judgement and Estimation** → 4. Reporting an Answer

Step 3: Estimation/Judgement



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

Step 4: Response



Understanding Measurement Error

- **Measurement Error** – occurs when the response provided by a respondent differs from the real or true value.

- Errors may be **random** or **systematic**
  - Random Error – caused by factors that randomly affect measurement of the variable across the sample. E.g. mood.
  - Systematic Error – not determined by chance but is an inaccuracy inherent in the whole sample

Sources of Measurement Error

- Questionnaire/Survey Instrument
- Respondent Bias
- Surveyor/Interviewer Bias
- Data Collection method

Questionnaire/Survey Instrument

Question Wording

- Ambiguous wording
  - *Is your work made more difficult because you are expecting a baby?*
- Complex question
  - *Has it happened to you that over a long period of time, when you neither practiced abstinence, nor used birth control, you did not conceive?*
- Abstract Concepts: *empowerment, bargaining power, social cohesion, perceptions and attitudes*
  - *How effective is your leader?*

Questionnaire/Survey Instrument

- Double-barreled question
  - *Do you agree that AIDS can be transmitted by shaking hands with a person with AIDS or through other means of physical contact?*
- Technical jargon
  - *What was your age when your menstrual period first started?*
- Leading questions
  - *Don't you agree that vegetarian food is tastier than non-vegetarian food?*

Questionnaire/Survey Instrument

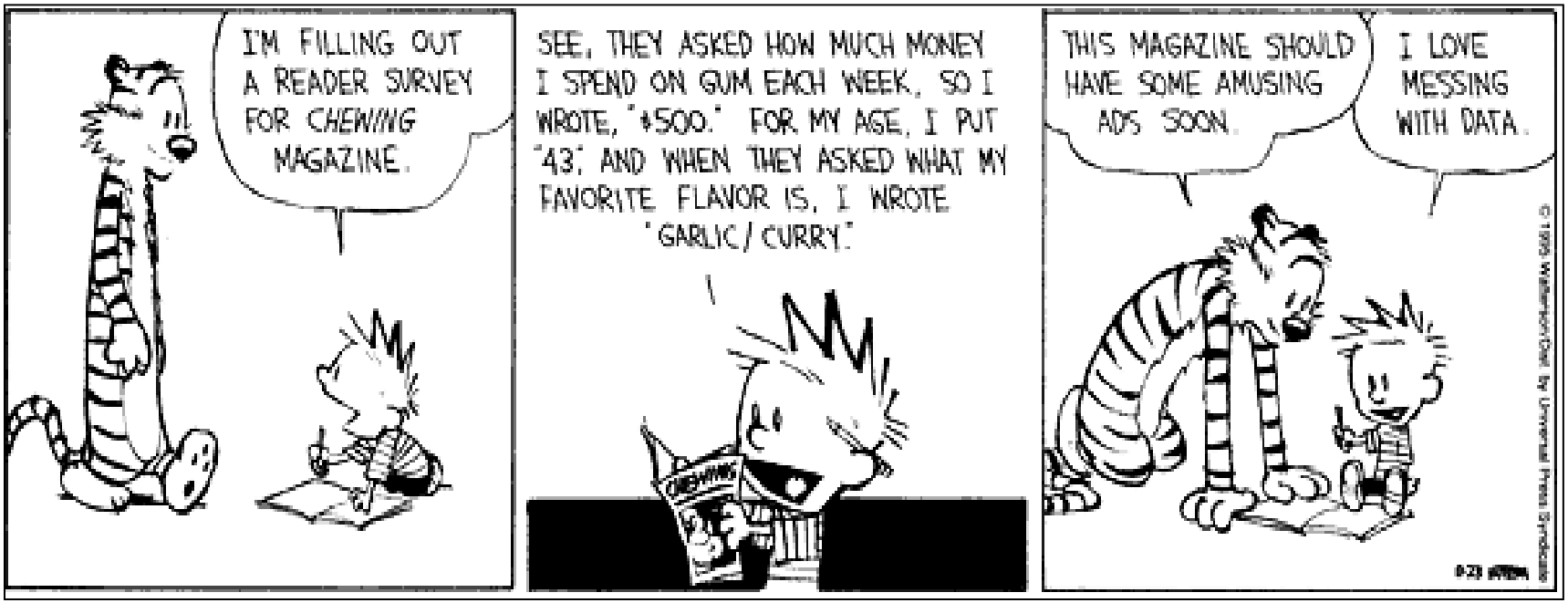
- Completeness
  - The response categories do not include all categories that can be expected as a response

    Q. What is the highest level of education completed?

    - Basic Education (1-5th)
    - Middle School (6th-8th)
    - High School (9th-12th)
    - College Degree
    - Post Graduate
    - Other Professional Degree (e.g. Medical, Law, Teacher)
- Overlapping categories

Respondent Bias

- **Response Bias**
  - General term for many types of cognitive/psychological biases



Respondent Bias – Recall Bias

- Differences in the accuracy or completeness of the recollections retrieved by respondents regarding events or experiences from the past
  - Example: distance to health center, profit, consumption, income etc.
- Depends on length of recall period, salience of past event to be recalled, respondent's motivation
  - Other factors, too many questions that require?

Respondent Bias – Mitigating Recall Bias

- Avoid mental calculations
- Consistency checks and shorter intervals
  - How much did you spend in the last week on x? How much did you spend in the last 4 weeks on x?
- Multiple measurements of same indicator
  - How many minutes does it take to walk to the health center? How many kilometers away is the health center?

Respondent Bias – Anchoring Bias

- Human tendency to accept and rely on the first piece of information received before making a decision. That first question can act as an anchor which sets the tone for subsequent questions.
  - Ex. Respondents asked 2 questions:
    1. Should professors be expelled for plagiarism?
    2. Should students be expelled for plagiarism?

- Review the structure and flow of questionnaire

Respondent Bias – Social Desirability Bias

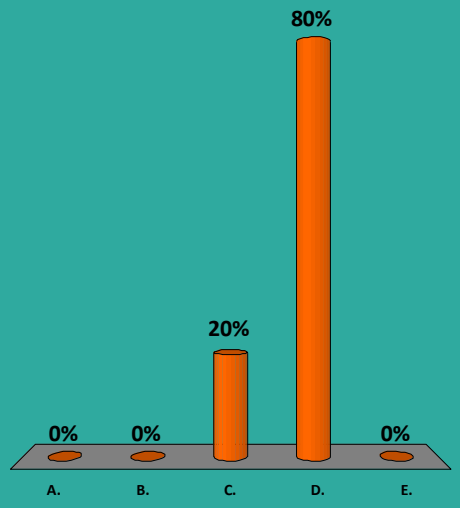
- Human tendency to respond with a more favorable answer, when respondent feels that their true answer would not be acceptable to the interviewer.
  - Tendency to 'over-report' good behavior and 'under-report' bad behavior
  - Illegal behavior, socially stigmatizing behavior
    - Ex. How often do you wash your hands after using the latrine? (Always, Often, Sometimes, Rarely, Never)

Where could the following question first produce error?

A. Validity
B. Comprehension
C. Retrieval
D. Judgment/ Estimation
E. Response

Q. How would you rate the quality of coffee this morning?

- Very good
- Somewhat good
- Not good

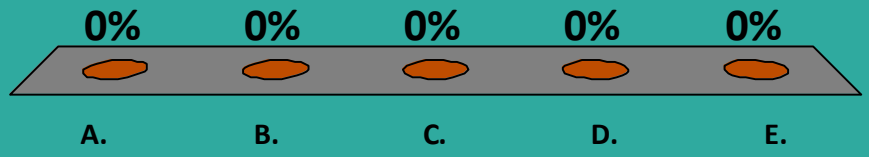


Where could the following question first produce error?

A. Validity
B. Comprehension
C. Retrieval
D. Judgment/ Estimation
E. Response

Q. Two new treatments have been developed to treat 600 terminally ill patients. Treatment A will save 200 people, while Treatment B will allow 400 people to die. Which treatment would you prefer?

- Treatment A
- Treatment B

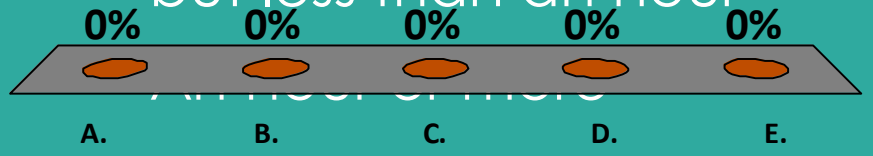


Where could the following question first produce error?

A. Validity
B. Comprehension
C. Retrieval
D. Judgment/ Estimation
E. Response

Q. In Arizona, some voters reported having to wait more than 5 hours to vote. How long did you have to wait last time you voted?

- No time (there was no line, or I voted by mail)
- Less than 10 minutes
- Between 10 minutes and 30
- More than 30 minutes but less than an hour



Surveyor/Interviewer Bias

- Respondents answers influenced by interviewer's behavior, appearance, question delivery, reaction, etc.

Problem with Data Collection Method

- Calibration errors lead to inaccurate measurements
  - e.g. non-zeroed weight scale
- For health measurements, could consider recording the measure thrice

How to Measure it (Well)

Sources of Measurement

Best Practices- Tips

Data collection on people

- Surveys
- Exams, tests, etc.
- Games
- Vignettes
- Direct Observation
- Diaries/Logs
- Focus groups
- Interviews

Survey: Modes of Data Collection

- Paper-based
- Phone applications
- IVRS
- GPS
- Biometric device
- Online surveys
- SMS alert system
- Smart cards

Choosing the timing of your survey

- There are several reasons why you might want to administer your survey **before** a program is rolled-out.
  - To identify needs
  - To show "control" and "treatment" groups are similar
  - To describe your population/sample of interest
- You might want to administer it **during** a roll-out.
  - To track implementation
  - To identify potential mechanisms of impact
- You might also want to administer it **after** a roll-out.
  - To evaluate program impact (fadeout and long-term effects)

Concluding Thoughts

Concluding Thoughts

- The process of collecting "good" data requires a lot of efforts and thought
- Theory of change guides measurement
- Indicators need to accurately measure
- 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

J-PAL
ABDUL LATIF JAMEEL POVERTY ACTION LAB

Thank you

Harini Kannan
harini.kannan@ifmr.ac.in

IFMR clear