



Measurement

Sara Heller



Course Overview

1. What is Evaluation?
2. Measurement
3. Why Randomize?
4. How to Randomize?
5. Sampling and Sample Size
6. Threats and Analysis
7. Start to Finish
8. Generalizability



J-PAL

ABDUL LATIF JAMEEL POVERTY ACTION LAB

Measurement

Outcomes, Indicators and Data



Lecture Overview

- What to measure
 - Theory of change
 - Purpose of measurement
- How to measure
 - Sources of measurement
 - Measurement concepts
 - Indicators and Indices
 - Response process
 - Challenges to measurement
 - Primary data
 - Administrative data
 - Ethics and IRB

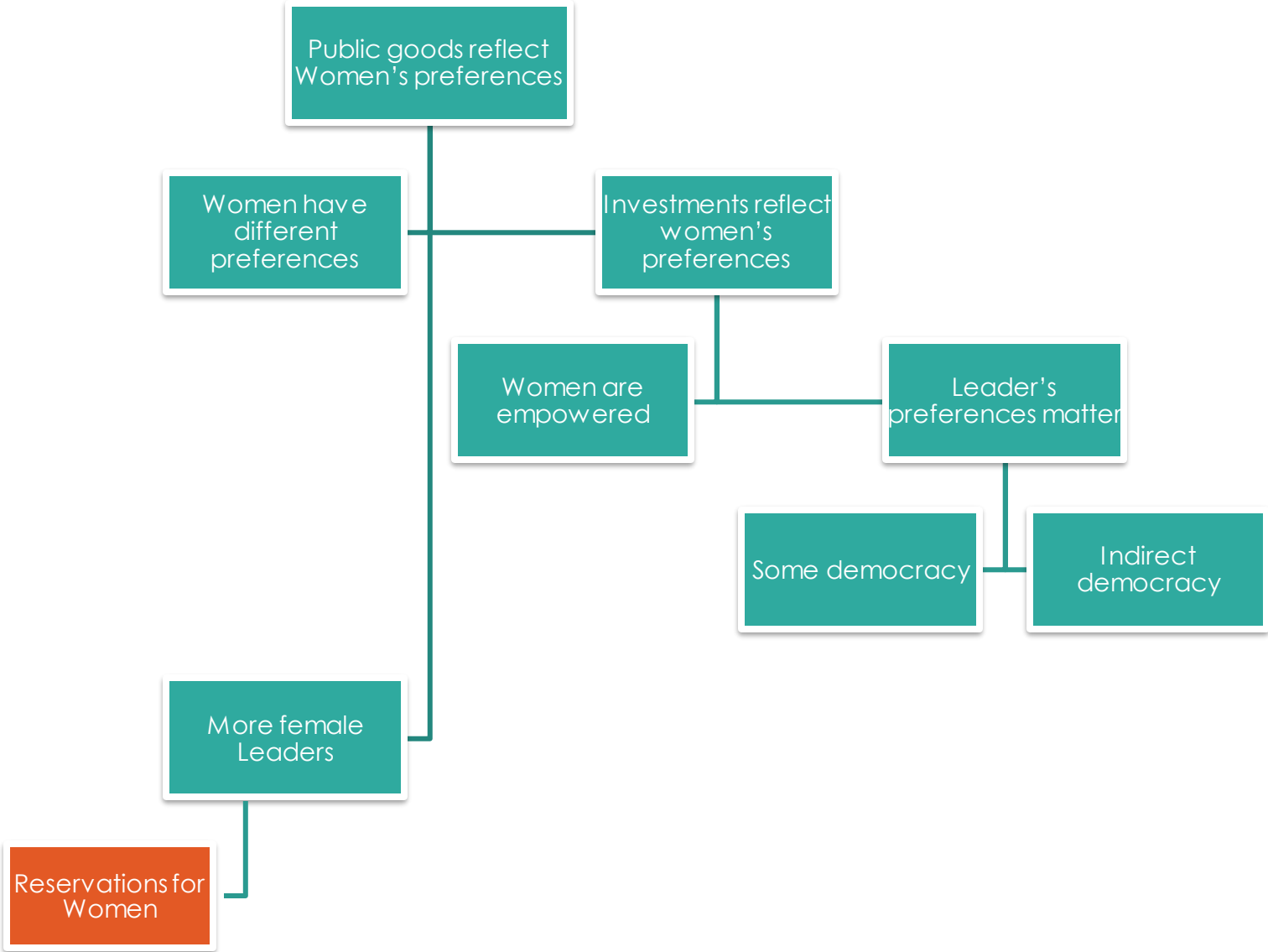
Lecture Overview

- What to measure
 - Theory of change
 - Purpose of measurement
- How to measure
 - Sources of measurement
 - Measurement concepts
 - Indicators and Indices
 - Response process
 - Challenges to measurement
 - Primary data
 - Administrative data
 - Ethics and IRB

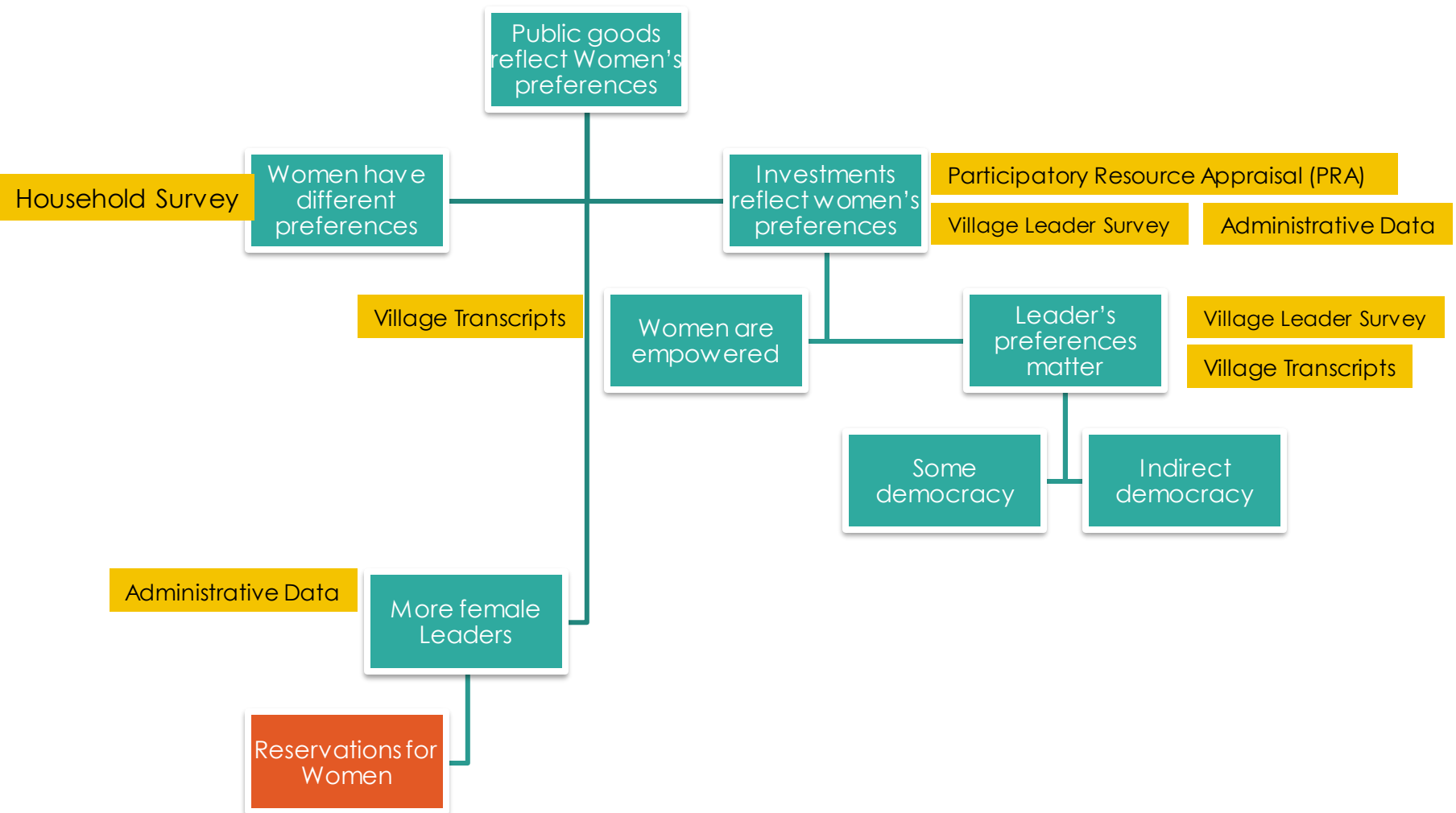
Lecture Overview

- What to measure
 - Theory of change
 - Purpose of measurement
- How to measure
 - Sources of measurement
 - Measurement concepts
 - Indicators and Indices
 - Response process
 - Challenges to measurement
 - Primary data
 - Administrative data
 - Ethics and IRB

Theory of Change: Women as Policymakers



Theory of Change: How to measure?



Log Frame

	Objectives Hierarchy	Indicators	Sources of Verification	Assumptions / Threats
Impact (Goal/ Overall objective)	Public good investment represents women's preferences	Government spending	Administrative data: Budgets, Balance Sheets	Pradhan preferences matter: imperfect/some democracy
Outcome (Project Objective)	Women voice political views	Number of times a woman spoke	Transcript from village meeting	Women develop independent views
Outputs	More female Pradhans	Whether or not a Panchayat had a female Pradhan	Administrative records	The law is implemented, there is no backlash
Inputs (Activities)	Reservations for women	Law is passed	The constitution	The government realizes the need for women representation

Source: Roduner, Schlappi (2008) Logical Framework Approach and Outcome Mapping, A constructive Attempt of Synthesis,

Results, By State, By Issue

		West Bengal			Rajasthan		
Issue	Investment Indicator	Issue Priority for		Investment Measure in Quota Villages	Issue Priority for		Investment Measure in Quota Villages
		W	M		W	M	
Drinking Water	# facilities	31%	17%	9.09*	54%	49%	2.62*
Road Improvement	Road Condition (0-1)	31%	25%	0.18*	13%	23%	-0.08*
Irrigation	# facilities	4%	20%	-0.38	2%	4%	-0.02
Education	Informal education center	6%	12%	-0.06	5%	13%	-

Lecture Overview

- What to measure
 - Theory of change
 - Purpose of measurement
- How to measure
 - Sources of measurement
 - Measurement concepts
 - Indicators and Indices
 - Response process
 - Challenges to measurement
 - Primary data
 - Administrative data
 - Ethics and IRB

Purpose of measurement

- Measure outcomes
 - Long term outcomes
 - Intermediate outcomes
 - First order outcomes
 - Second order outcomes
 - Inputs, outputs, etc
- Treatment compliance (individual and group level)
 - Predictors of compliance
- Covariates
- Heterogeneous treatment effects
- Context for external validity

Lecture Overview

- What to measure
 - Theory of change
 - Purpose of measurement
- How to measure
 - Sources of measurement
 - Measurement concepts
 - Indicators and Indices
 - Response process
 - Challenges to measurement
 - Primary data
 - Administrative data
 - Ethics and IRB

Lecture Overview

- What to measure
 - Theory of change
 - Purpose of measurement
- How to measure
 - Sources of measurement
 - Measurement concepts
 - Indicators and Indices
 - Response process
 - Challenges to measurement
 - Primary data
 - Administrative data
 - Ethics and IRB

First-order questions in measurement

- What data do you collect?
- Where do you get it?
- When do you get it?

Where can we get data?

- Obtained from other sources
 - Publicly available
 - Administrative data
 - Other secondary data
- Collected by researchers
 - Primary data
 - Survey and non-survey methods

PLACE OF BIRTH		SEX	AGE	RACE	MARITAL STATUS	EDUCATION	INDUSTRY	RELIGION	REMARKS
1	John ...	M
2	Mary ...	F
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50



https://commons.wikimedia.org/wiki/File:Cuyahoga_County_US_Census_Form-Herbert_Birch_Kingston_1920.jpg
[https://commons.wikimedia.org/wiki/File:US_Navy_020123-N-97602-004_Hospital_Corpsman_2nd_Class_Jennifer_Ross_files_medical_records_ aboard_the_aircraft_carrier_USS_Nimitz_\(CVN_68\).jpg](https://commons.wikimedia.org/wiki/File:US_Navy_020123-N-97602-004_Hospital_Corpsman_2nd_Class_Jennifer_Ross_files_medical_records_ aboard_the_aircraft_carrier_USS_Nimitz_(CVN_68).jpg)

Types and Sources of Data

Information about a person/ household / possessions

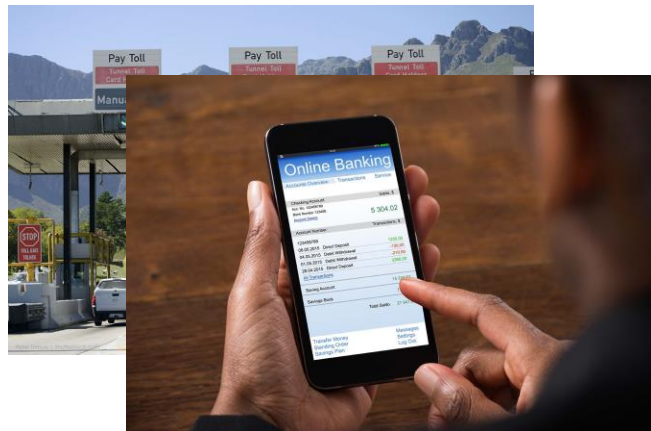
NOT about a person/ household / possessions

Information provided by a person



Objective and subjective measures

Automatically generated



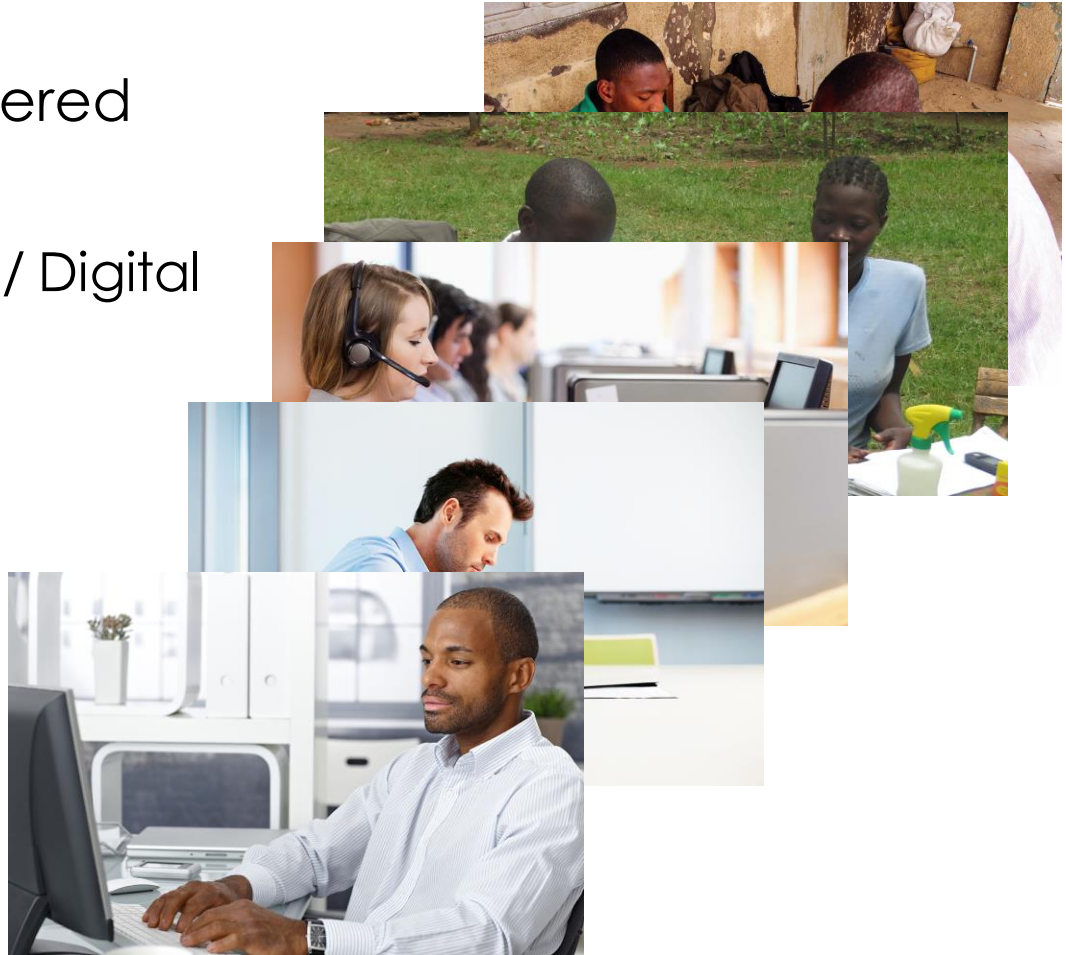
Primary Data Collection

- 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

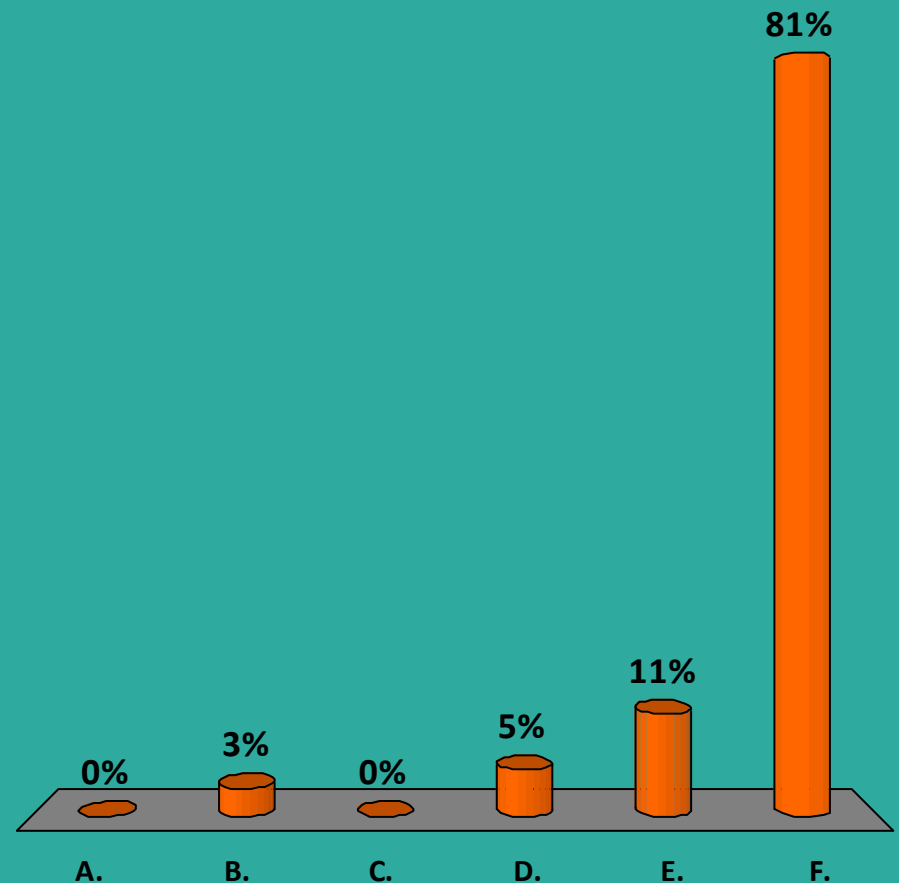


Primary Data: Who to collect data from?

- Respondent
 - Head of Household?
- Target respondent: should be most informed person for each module. Respondents for each module could vary.
- Survey setting

Which of these is an example of administrative data?

- A. Text from a tweet
- B. Information from a birth certificate
- C. Demographic information collected during a baseline survey
- D. Income information from tax records
- E. All of the above
- F. B and D



Administrative data

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

- Medical records
- Educational records
- Arrest records
- Banking records
- Personnel records



Administrative Data: Sources

- J-PAL's [Catalog](#) of US data sets
- Health
 - Vital statistics office
 - Health facilities (e.g., hospitals, clinics)
- Finance
 - Banks, credit unions
 - Credit rating agencies
- Education
 - Schools
 - Department of education
- Statistics department

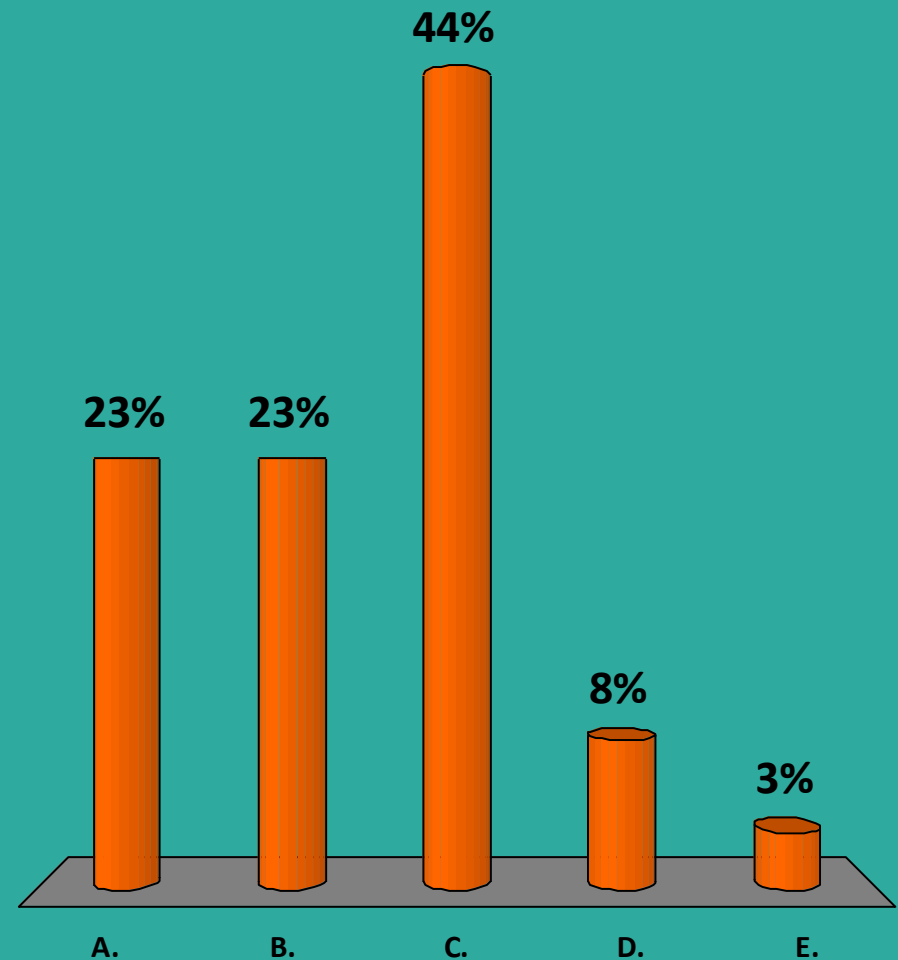
Why are administrative data useful?

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

- Available retrospectively
- Enable long-term follow-up
- Reduce logistical burden
- Include near census of relevant population
- Often cheaper than surveys
- May minimize bias and error

What source of data do you primarily expect to use in your work?

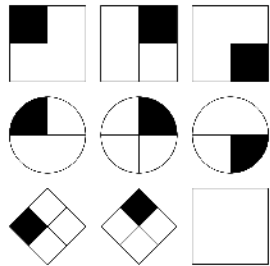
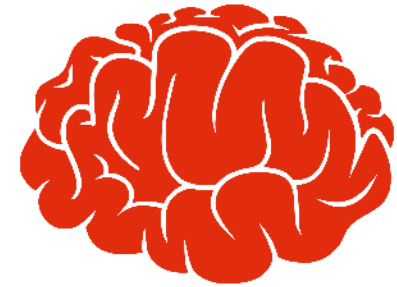
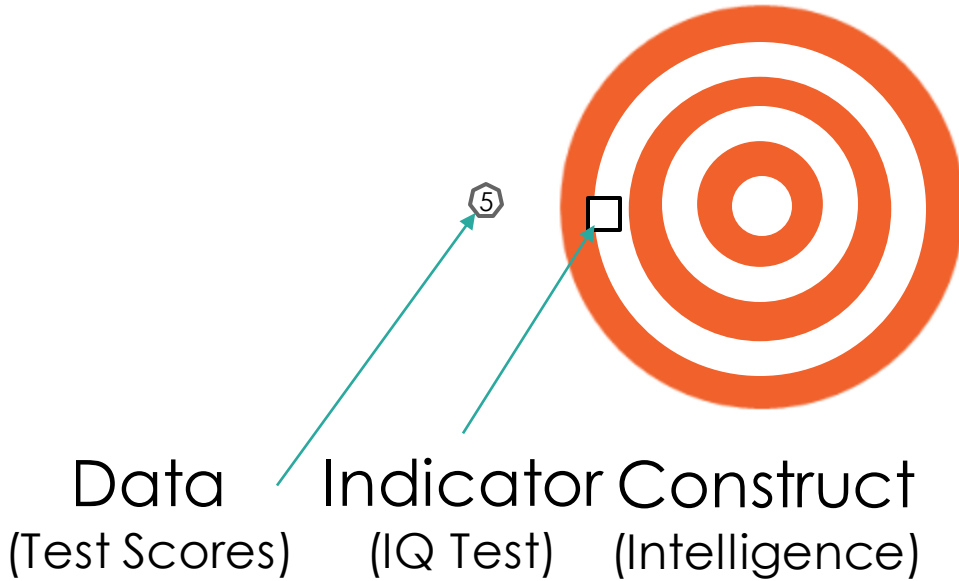
- A. Primary survey data
- B. Other primary (non-survey) data
- C. Administrative data
- D. Public data
- E. Other secondary data



Lecture Overview

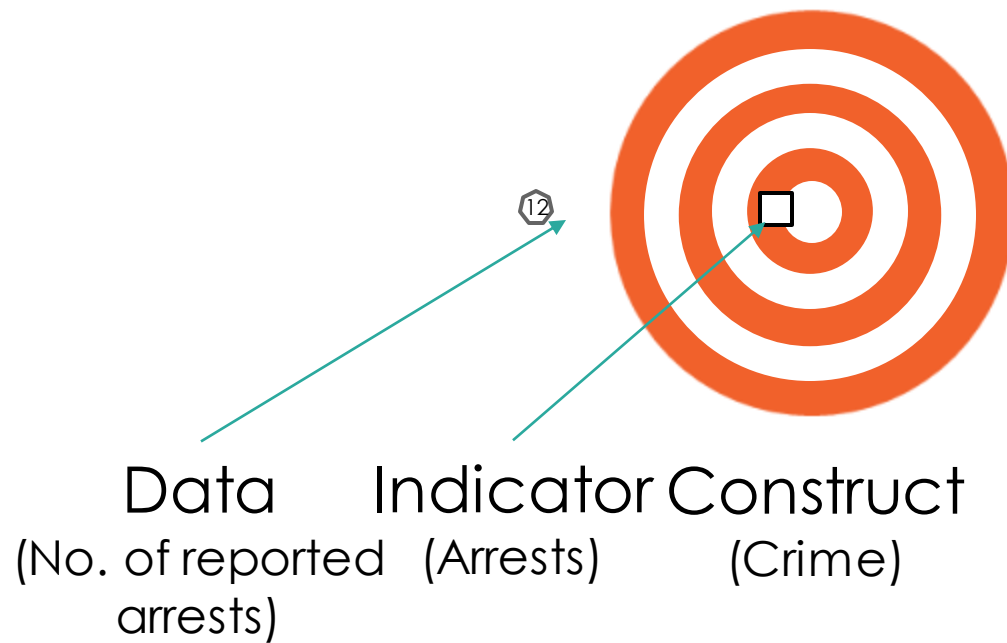
- What to measure
 - Theory of change
 - Purpose of measurement
- How to measure
 - Sources of measurement
 - Measurement concepts
 - Indicators and Indices
 - Response process
 - Challenges to measurement
 - Primary data
 - Administrative data
 - Ethics and IRB

Concept of measurement

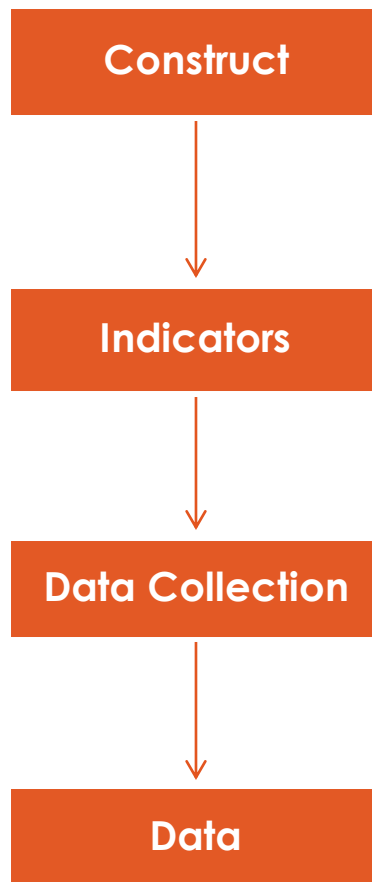


https://commons.wikimedia.org/wiki/File:Red_Silhouette_-_Brain.svg

Concept of measurement

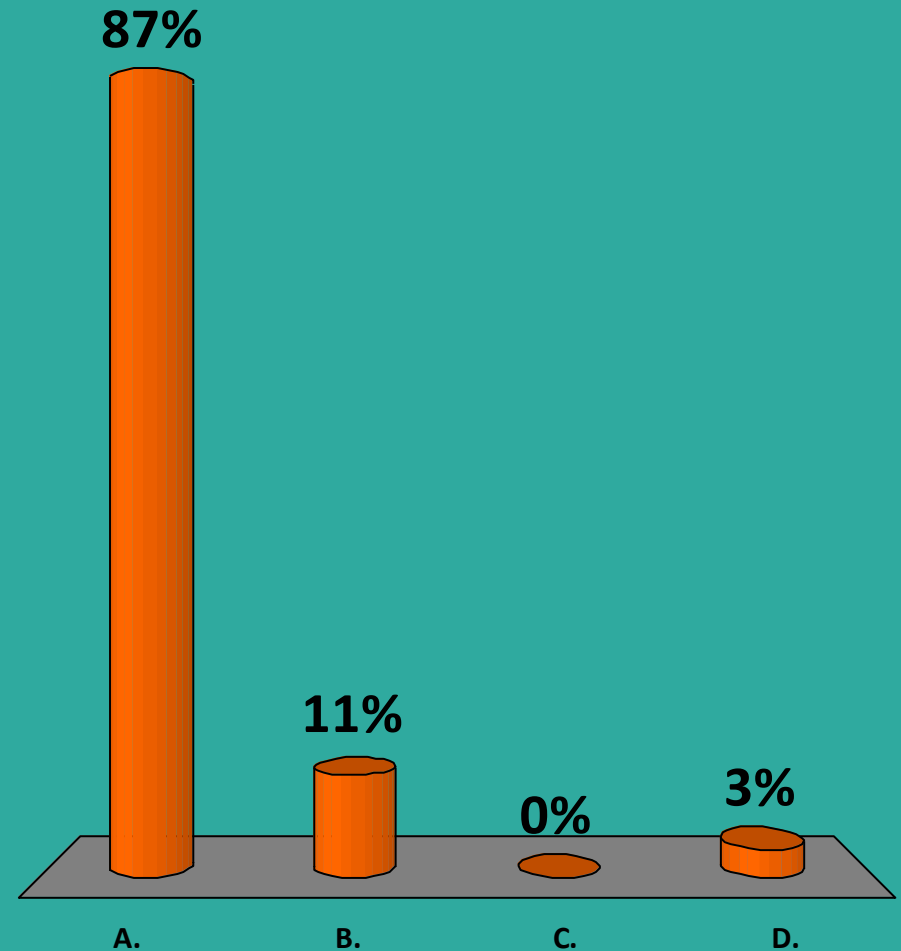


Concept of measurement



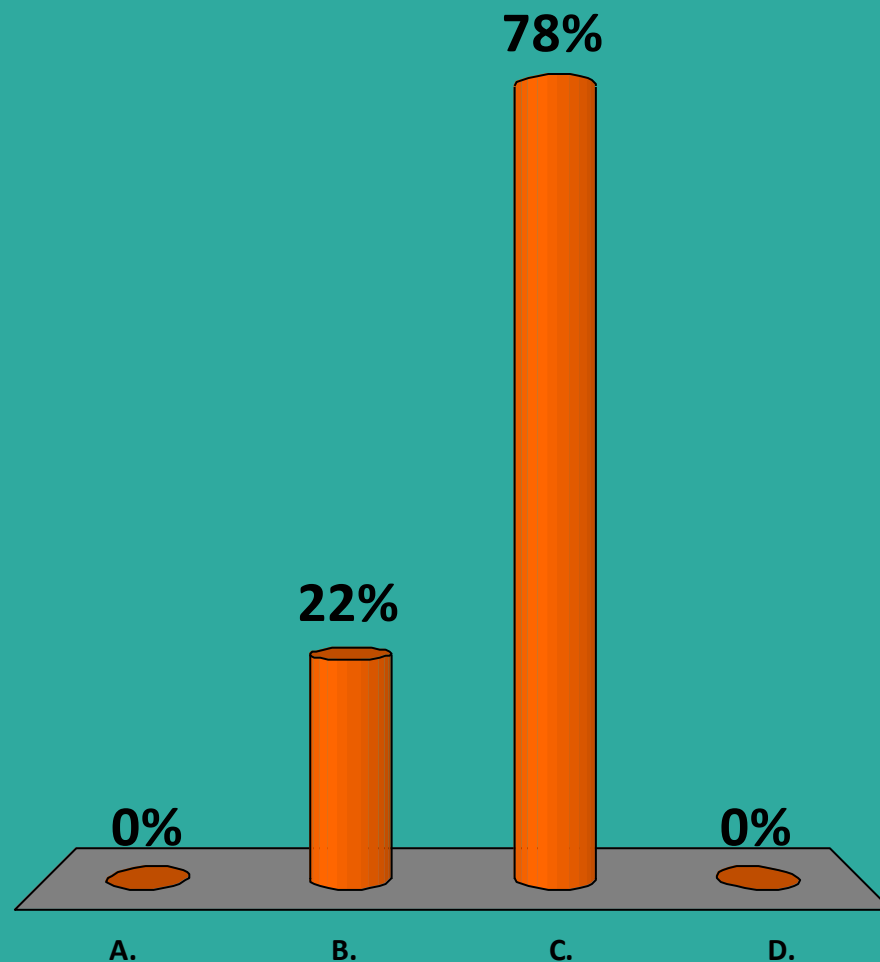
Empowerment is:

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



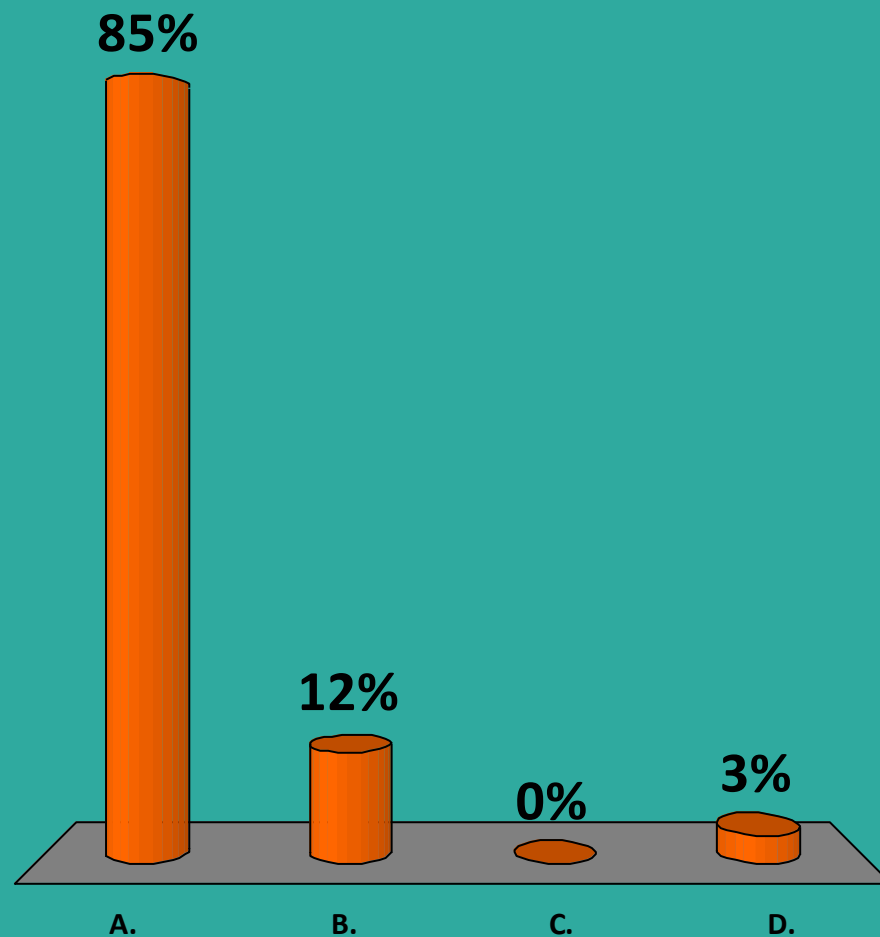
“Attendance = 170 days” as found in a School register is:

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



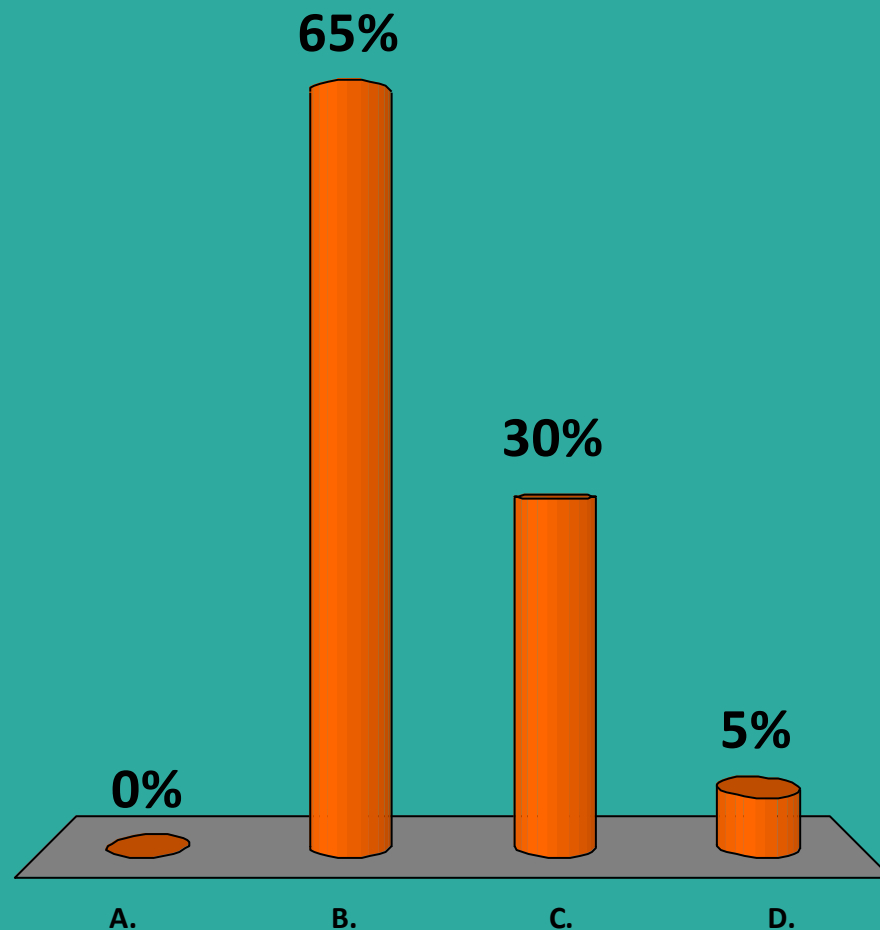
Discrimination is:

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



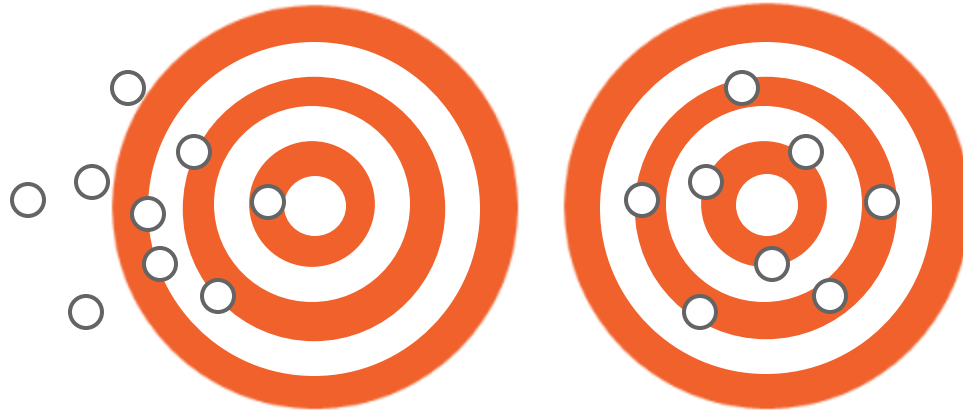
Hospital visits per month as found in an electronic medical record

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



The goals of measurement

- Accuracy
- Unbiasedness
- Validity

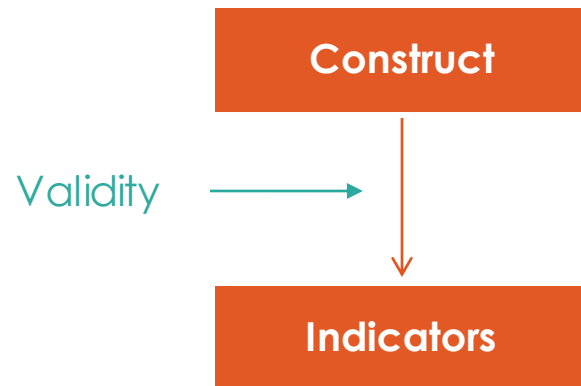


- Precision
- Reliability



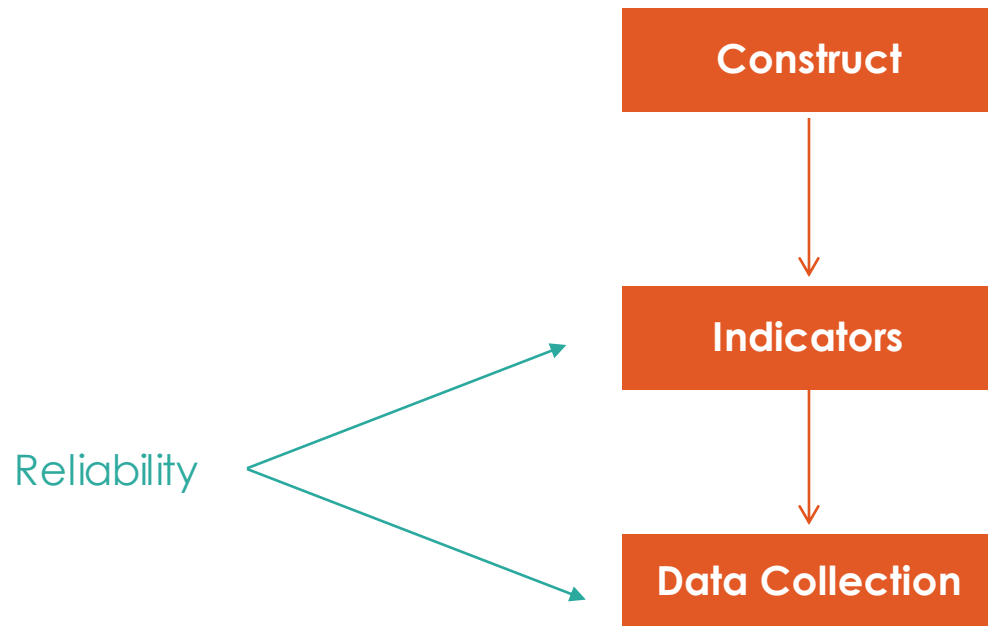
Validity

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



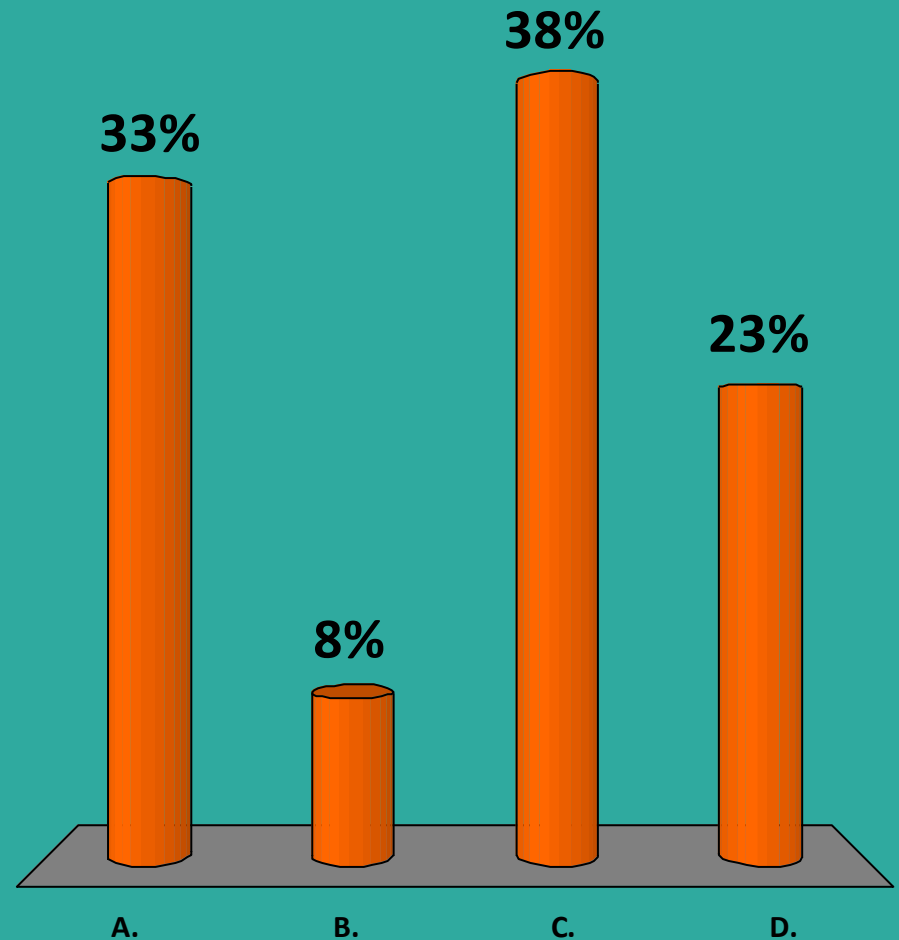
Reliability

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



Which is worse?

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



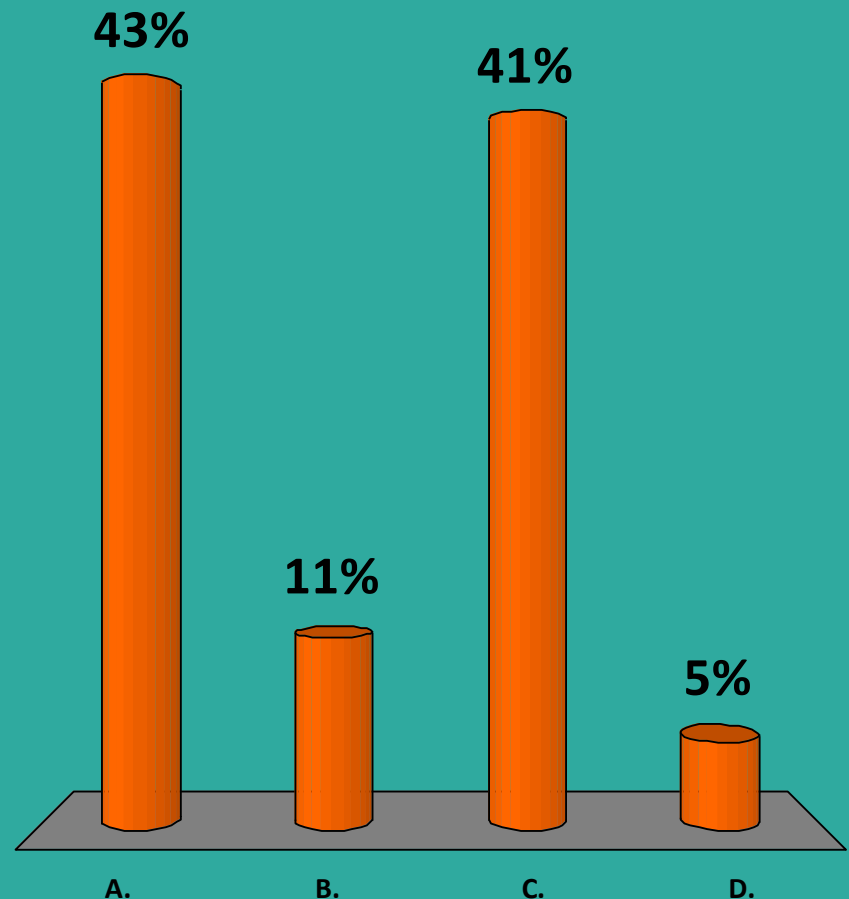
The problem

- With the following questions...

Outcome: well-being

Indicator: annual income

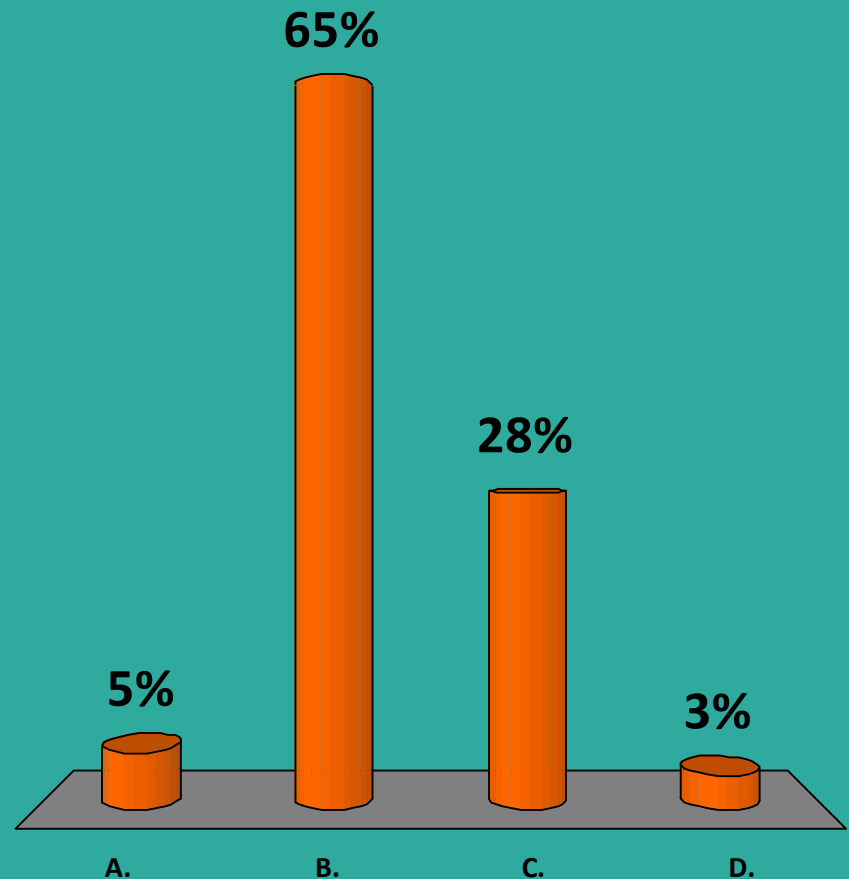
- A. Validity
- B. Reliability
- C. Both
- D. Neither



Outcome: consumption

Indicator: expenditure in last week

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

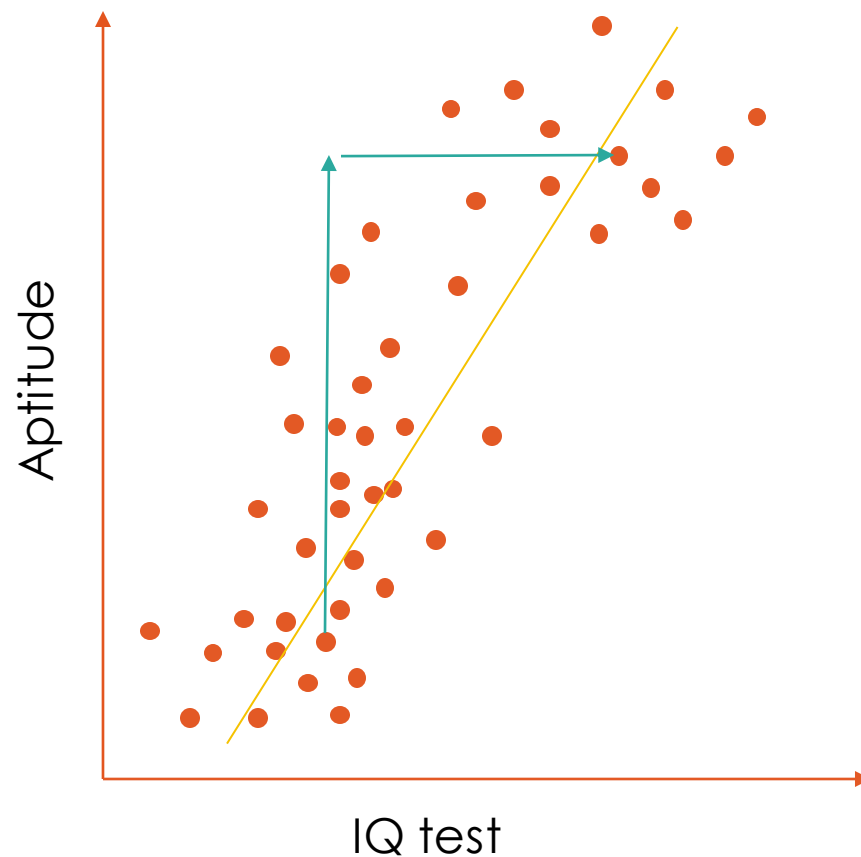


Lecture Overview

- What to measure
 - Theory of change
 - Purpose of measurement
- How to measure
 - Sources of measurement
 - Measurement concepts
 - Indicators and Indices
 - Response process
 - Challenges to measurement
 - Primary data
 - Administrative data
 - Ethics and IRB

Proxy indicators

- Constructs are hard to measure
 - Sensitive
 - Unknown
- Proxy indicators
 - Correlated with construct
 - Correlation is dynamic



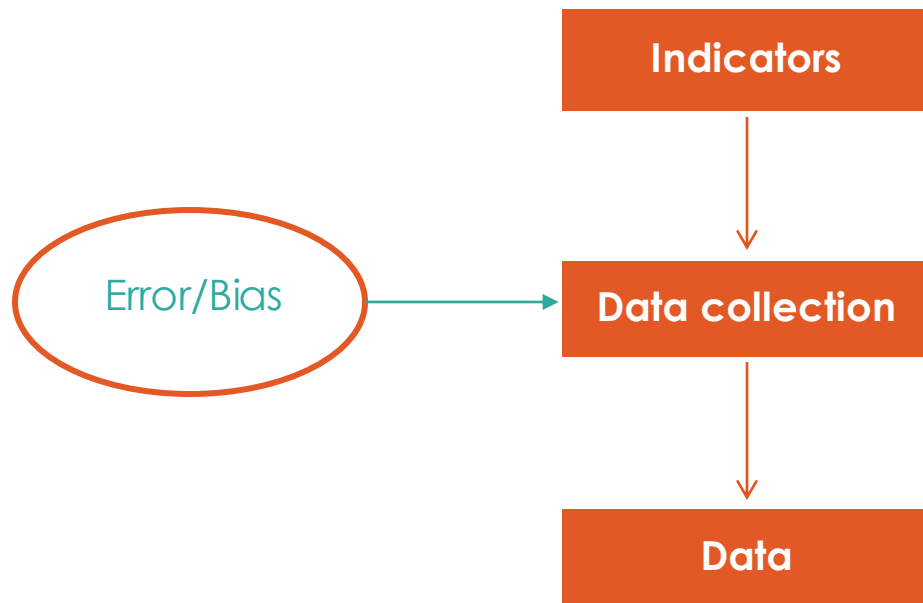
Indices: Examples

- Economy
- Prices
- Corruption index
- Test scores
- Women empowerment

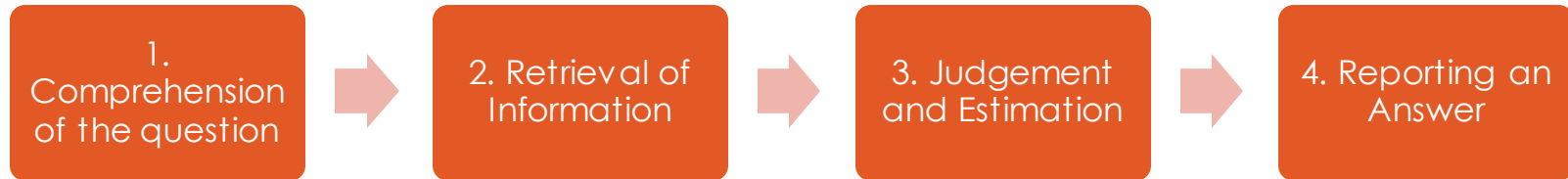
Lecture Overview

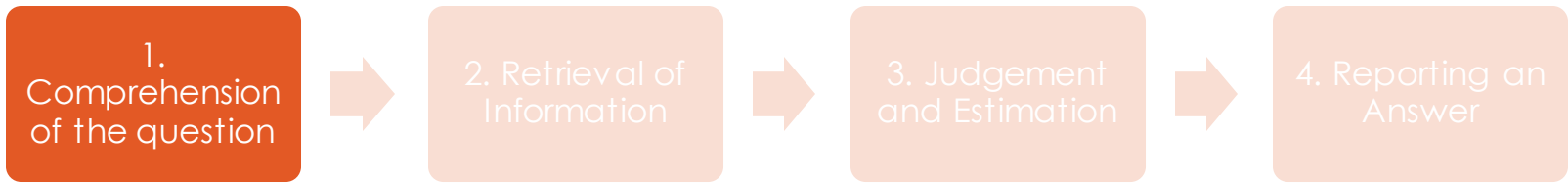
- What to measure
 - Theory of change
 - Purpose of measurement
- How to measure
 - Sources of measurement
 - Measurement concepts
 - Indicators and Indices
 - Response process
 - Challenges to measurement
 - Primary data
 - Administrative data
 - Ethics and IRB

The Response Process



Four-step Response Process

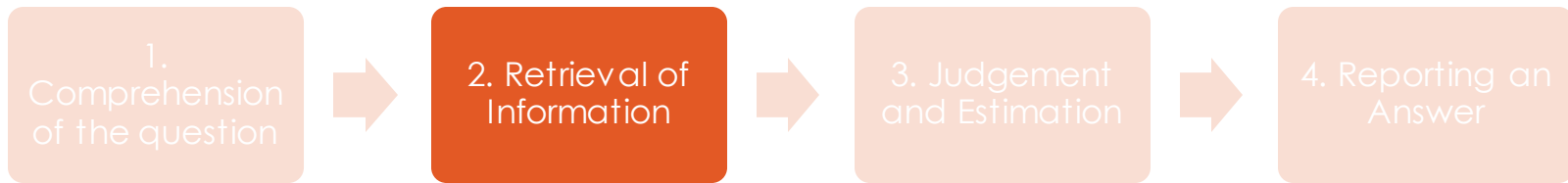




Step 1: Comprehension

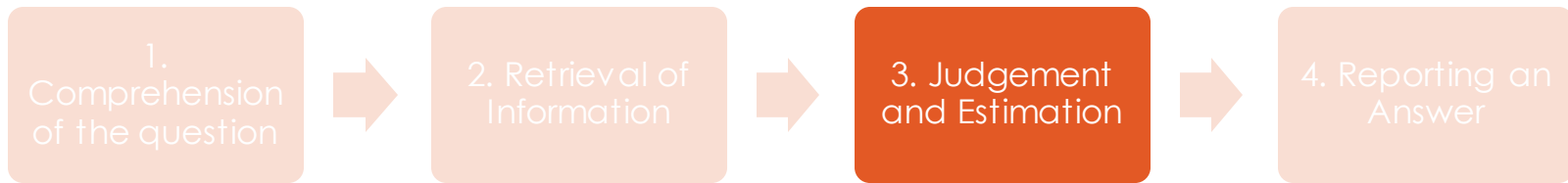
1.1 Total monthly income, before taxes





Step 2: Retrieval

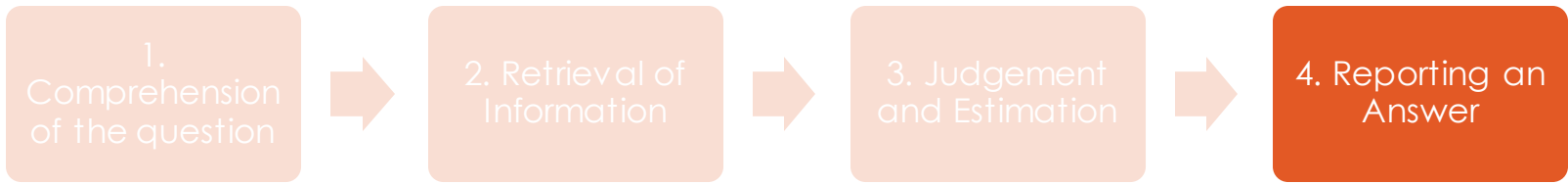




Step 3: Estimation/Judgement

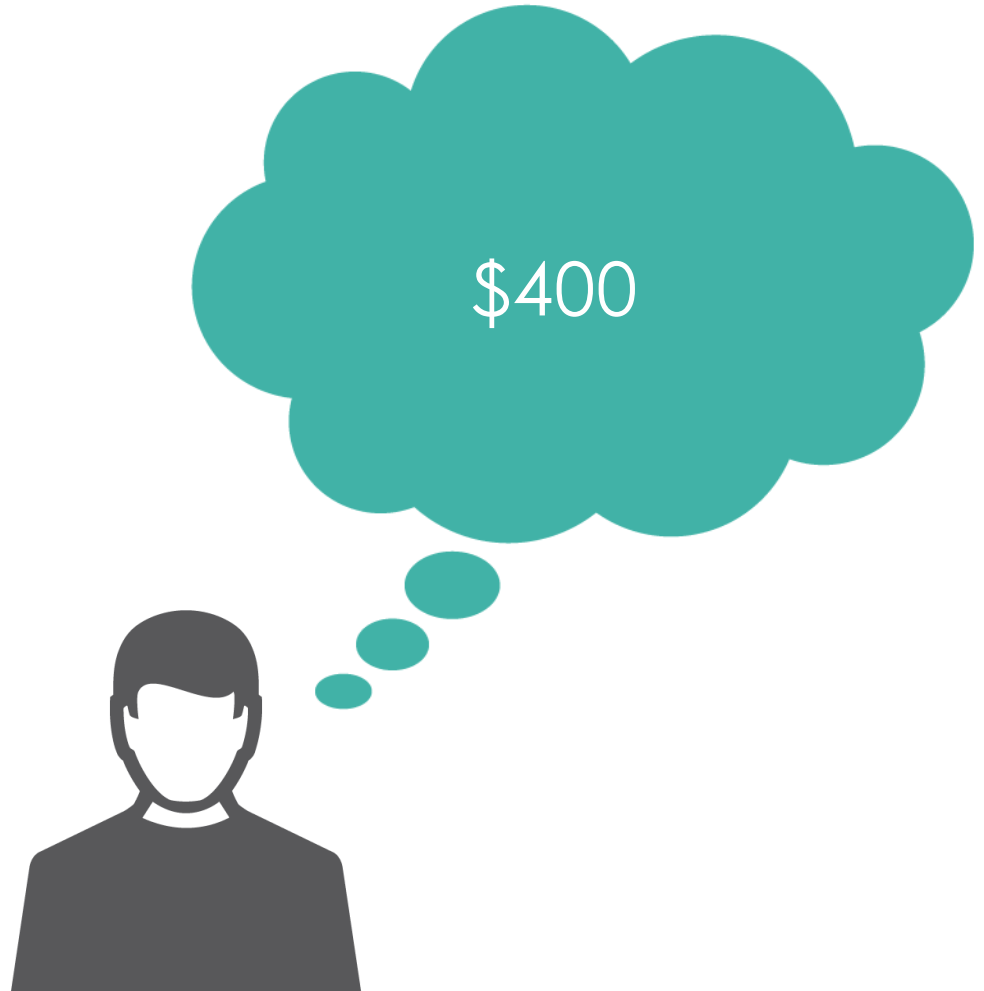
Social = \$ 200 per month
Workers' Compensation = 0
Pension = \$220 per month
What else??





Step 4: Response

1.1 Total monthly income, before taxes



SNAP Application Income Questions

First page

Total monthly income, for you and anyone who is applying, before taxes are taken out: \$ _____

Second page

Income

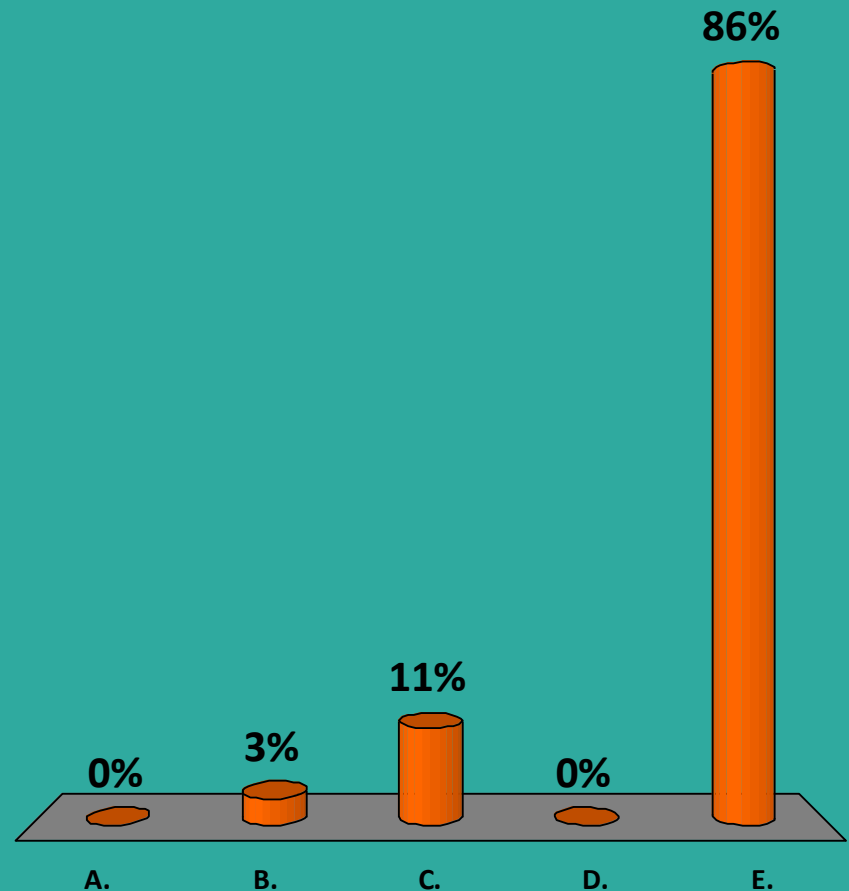
Does anyone have any income? Yes No *If yes*, list income you have already received this month or expect to receive this month.

Income includes, but is not limited to: Social Security; Pensions; SSI; Child Support; Alimony; Unemployment or Workers' Compensation; Pensions; Dividends or Interest; Room and Board; Private disability insurance; Veteran's Benefits; IRA Distributions and Annuity Payments.

PERSON WITH INCOME	TYPE/SOURCE OF INCOME	HOW MUCH	HOW OFTEN	DATE RECEIVED
		\$		
		\$		
		\$		
		\$		
		\$		

Which stage in the response process might produce measurement error?

- A. Comprehension
- B. Retrieval
- C. Estimation/Judgment
- D. Response
- E. All of the above



Lecture Overview

- What to measure
 - Theory of change
 - Purpose of measurement
- How to measure
 - Sources of measurement
 - Measurement concepts
 - Indicators and Indices
 - Response process
 - Challenges to measurement
 - Primary data
 - Administrative data
 - Ethics and IRB

Lecture Overview

- What to measure
 - Theory of change
 - Purpose of measurement
- How to measure
 - Sources of measurement
 - Measurement concepts
 - Indicators and Indices
 - Response process
 - Challenges to measurement
 - Primary data
 - Administrative data
 - Ethics and IRB

Measurement Error: Social Desirability Bias

Example:

Q. Were you arrested in the past month?

- Yes
- No

Tendency of respondents to answer questions in a manner that is favorable to others i.e., emphasize strengths, hide flaws, or avoid stigma

Respondents would be shy to admit to such behavior

Ask indirectly, ensure privacy

Measurement Error: Context effects

Framing within questionnaire

Example:

Q1. How many years of education do you have?

Q2. Did you go to public or private school?

Q3. Did your school provide everyone a quality education?

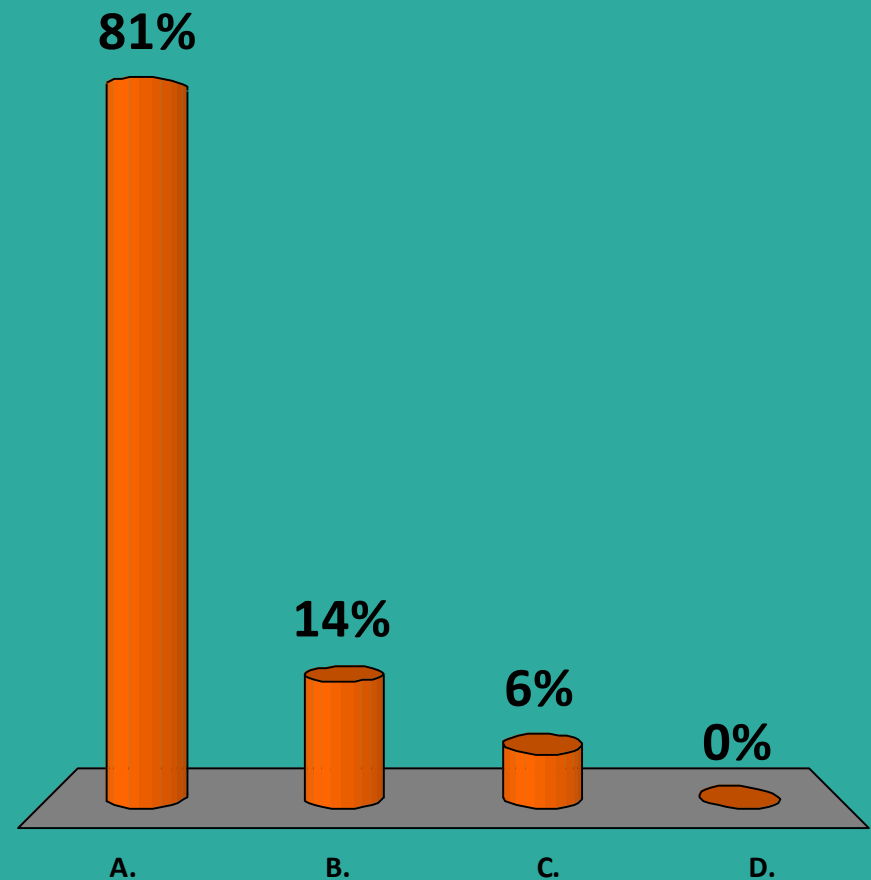
Q4. For the upcoming election, what are the top policy priorities you are looking for from the candidates?

Suddenly, education becomes everyone's top priority

Be careful of where questions are placed

How many meals have you eaten in the past hour?

- A. 0
- B. 1
- C. 2
- D. 3



Measurement Error: Telescoping Bias

Example:

Q. Did you purchase a TV or other electronic (worth over \$500) in the past 12 months?

People perceive recent events as being more remote than they are (backward telescoping) and distant events as being more recent than they are (forward telescoping)

This will lead to over reporting due to forward telescoping of events that happened before 12 months ago

Visit once at the beginning of the reference period, ask the question. Then ask, “since the last time I visited you, have you...?”

Other things to consider

- Question wording
 - Specific and easy to understand
 - Avoid negatives, double barreled questions etc.
- Translation
 - Back-translate and pretest in local languages
- Surveyor training/quality
- Mode of data collection/Data entry
- Survey length and respondent fatigue
- Piloting

Lecture Overview

- What to measure
 - Theory of change
 - Purpose of measurement
- How to measure
 - Sources of measurement
 - Measurement concepts
 - Indicators and Indices
 - Response process
 - Challenges to measurement
 - Primary data
 - Administrative data
 - Ethics and IRB

Access to Administrative Data: Generating an electronic file

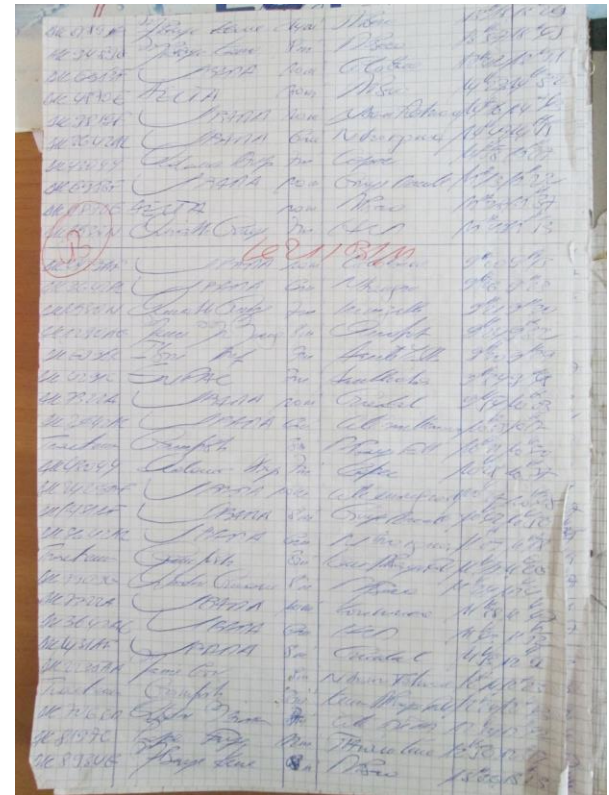
Records are in an unusable format

- Hand-written records
- PDF file

To address records in an unusable format

Digitize

- Housing records in a juvenile detention center



Access to Administrative Data: Getting access to an electronic file

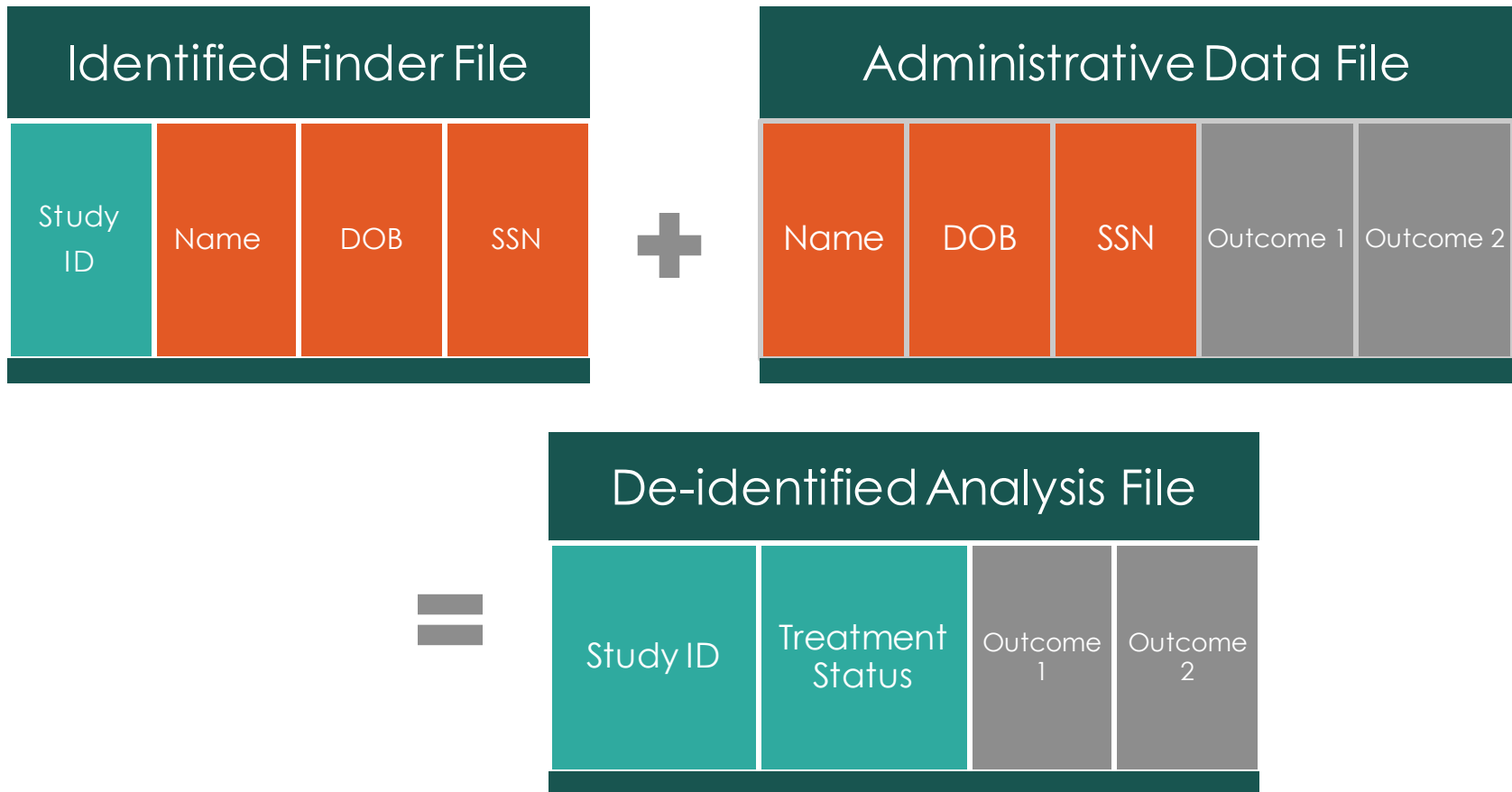
Regulations that limit access to identified data

- **US: Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule**
 - Obligations depend on level of identification and your contract with the data provider
 - Data security requirements
 - Fines for data leaks
 - Individual Authorizations or waivers (similar to informed consent)
- **US: Family Educational Rights and Privacy Act (FERPA)**
 - Protects students' educational records
 - Requires student consent for release of records
- **Other regulations depending on country context**

Using Administrative Data: Matching

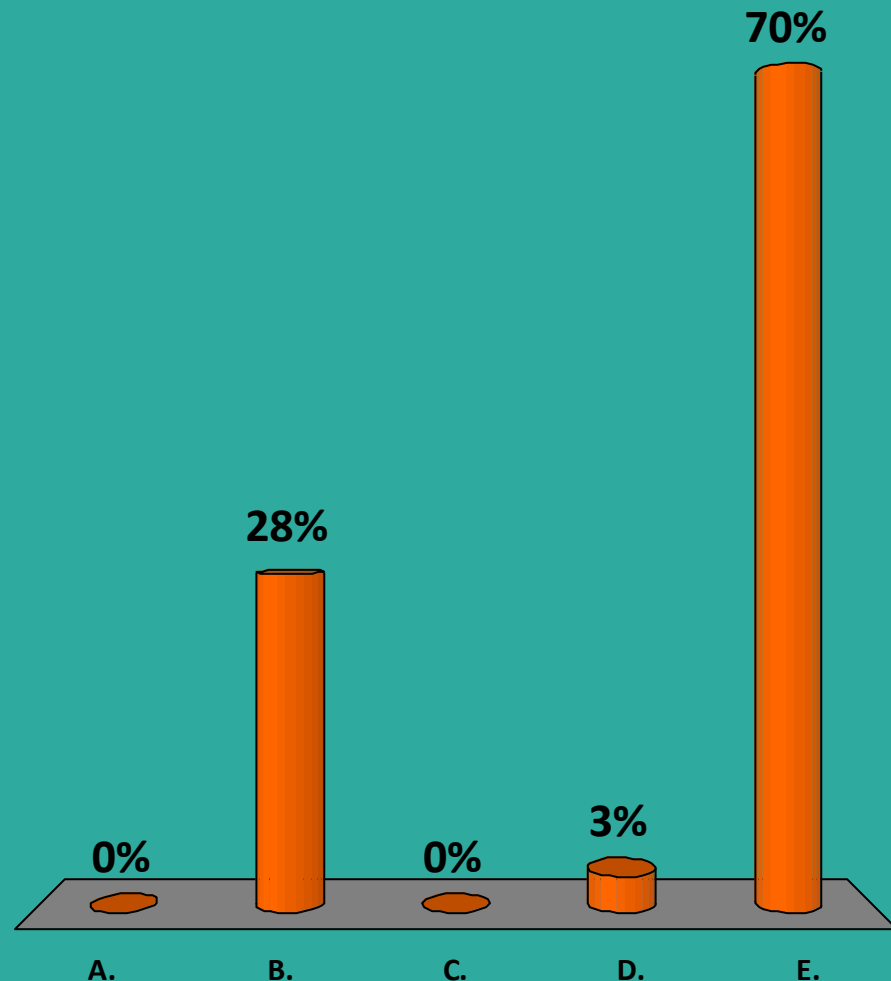
- **Personally Identifiable Information (PII)**
 - Name, Identification numbers, Address
 - Photos or biometric characteristics
- **Understand what identifiers the data agency collects**
 - Collect those same identifiers from your study sample at baseline
- **Use numeric identifiers instead of string variables**
 - Identification number and DOB instead of names and addresses
- **Participants may not be willing to provide sensitive identifiers**
 - E.g., identification numbers. Emphasize privacy & confidentiality during study enrollment

Separate identifiers from outcomes



When choosing identifiers for matching study data to administrative data, which of the following identifiers would be preferable to using an individual's street address?

- A. An email address
- B. A government-issued, unique identification number
- C. Date of birth
- D. All of the above
- E. B and C



Determining Data Accuracy

How can we ensure that the data are accurate?

- Unlike with survey data, the researcher does not have a say in the data collection and processing phase



Determining Data Accuracy

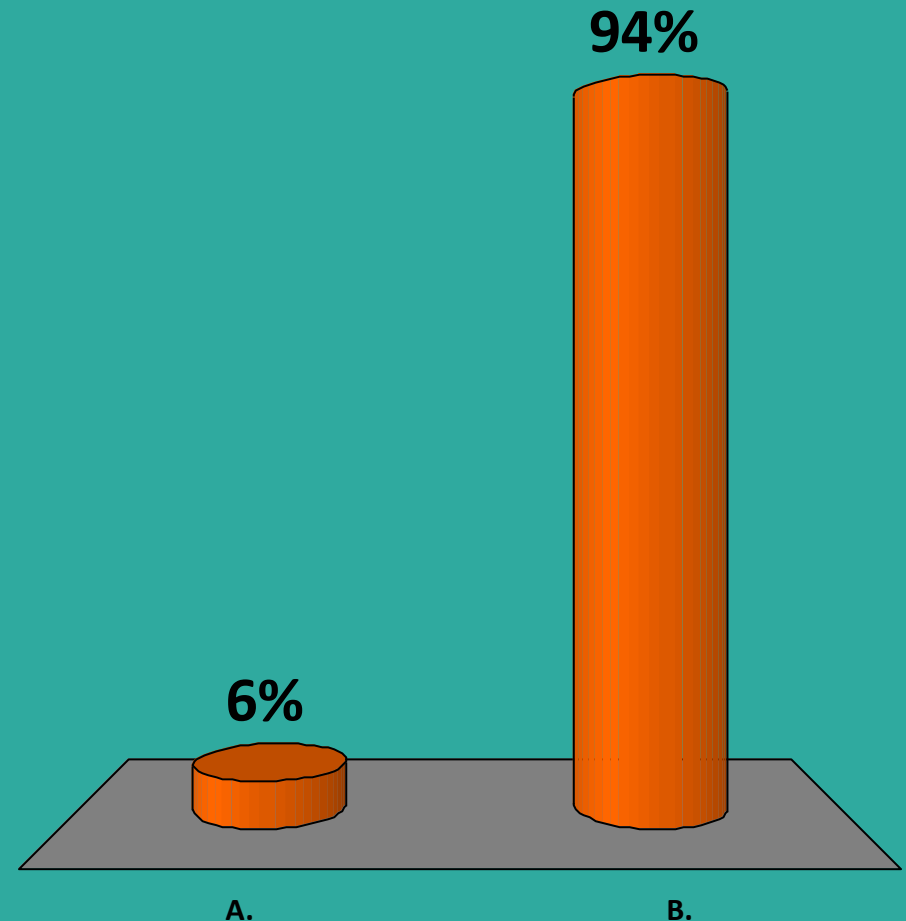
To address possibly inaccurate data

- **Cross-reference with other sources to ensure accuracy**
- **Identify the data agency's quality control protocol**
- **Choose indicators that are unlikely to be incorrectly reported**
 - Select variables that are straightforward and less susceptible to human error
 - Request raw variables
- **Communicate with program or implementing partner responsible for collecting data**
 - Ask how and why data are collected

Unlike survey data, administrative data are not susceptible to bias.

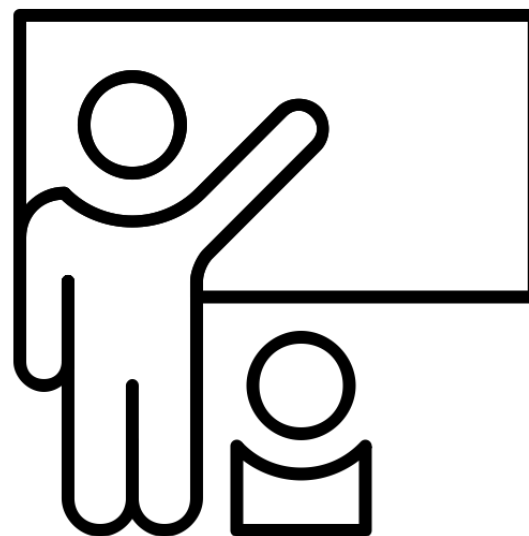
A. True

B. False



Reporting Bias

- From an individual
 - E.g., under-reporting income to qualify for a social welfare program
- From an administrative organization
 - E.g., schools over-report attendance to meet requirements



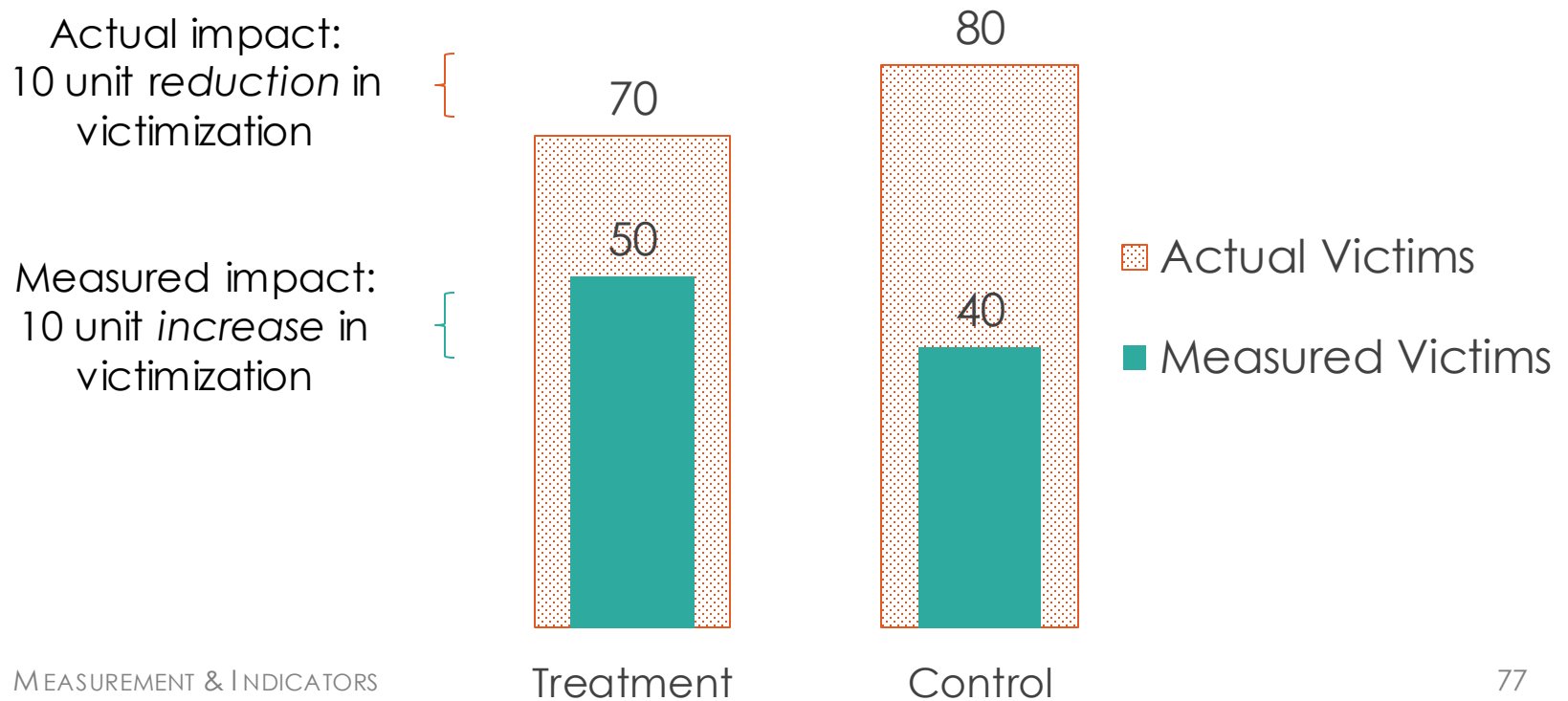
Reporting Bias

To address reporting bias

- **Identify the context in which the data were collected**
 - Were there incentives to misreport information?
- **Choose variables that are not susceptible to bias**
 - E.g., hospital visit v. value of insurance claim

Differential Coverage

- Differential ability to **link** individuals to administrative records
- Treatment and control are differentially likely to **appear** in administrative records
 - E.g., victimization as measured by calls to report crime



Differential Coverage

To address differential coverage bias

- **Collect identifiers for linking during the baseline survey**
 - To ensure that you are equally likely to be able to **link** treatment and control individuals to their records
- **Identify the data universe**
 - Which individuals are included in the data and which are excluded, and why?
 - To ensure the intervention does not affect the likelihood of **appearing** in a data set
- **Identify how the intervention may affect the reporting of outcomes**
 - Identify the context in which the data were collected
 - Determine direction in which estimates are likely to be biased

Lecture Overview

- What to measure
 - Theory of change
 - Purpose of measurement
- How to measure
 - Sources of measurement
 - Measurement concepts
 - Indicators and Indices
 - Response process
 - Challenges to measurement
 - Primary data
 - Administrative data
 - Ethics and IRB

Ethics and IRB

- “Experimenting on people”
- Belmont Principles
 - Respect for persons
 - Beneficence
 - Justice
- Institutional Review Boards (IRBs)
- Informed Consent
 - Consent to your use of their primary or administrative data
 - IRBs and/or data providers may require that individuals consent to **each specific data set** that may be used
 - Waiver of informed consent



END

