

# Surveys and the Research Process

Chase H. Harrison, Ph.D.  
Department of Government, FAS  
Program on Survey Research, IQSS  
Harvard University



# Research Design

---

# Two General Types of Research

## Discovery

- Conjecture,
- Exploration
- Innovation
- Theory formation

## Appraisal

- Assessment
- Evaluation
- Demonstration
- Justification
- Proof
- Testing
- Verification/falsification

# Inference in Social Sciences

”Goal is to make descriptive or explanatory **inferences about [about social and political life]** on the basis of empirical information about the world....That.. may involve ***descriptive inference*** – using observations about the world to learn unobserved facts. Or.. ***causal inference*** – learning about causal effects from the data observed. The **domain of inference may be restrictive in space and time** – voting behavior in American elections since 1960 or social movements in Eastern Europe since 1989, **or it can be extensive** – human behavior since the start of agriculture.”

King, Keohane, Verba (1994) *Designing Social Inquiry*



# Social Science Research

A photograph of two parrots perched on a wooden branch. The parrot on the left is in focus, showing its colorful feathers. The parrot on the right is out of focus. The background is a warm, yellowish light, possibly from a lamp.

Goal is to make descriptive or explanatory inferences about the world based on empirical observations

Using things that you know to learn things that you don't know

Using things that you can see to learn about things that you can't see

Source: King, Keohane and Verba: Designing Social Inquiry

# Scientific Social Science Research

Goal is to make descriptive or explanatory inferences about the world based on empirical observations

Procedures are public (tractable)

Findings are uncertain

Follows scientific methodology



# Examples:

- Populations

- Agents in different systems
- Decision influencers
- Voters
- Consumers of different product types
- Early-adopters
- Activists in Social Movements
- Civil Wars
- Large organization

- Topics of Inference:

- Hiring decisions
- Brand Value
- Intergroup relations
- Institutions to Facilitate Democratic Participation
- Voting in Highly Polarized Settings
- Programs to incentivize peace
- Etc.

# Construct Example: Intergroup Relations (e.g. Individual Attitudes toward Members of Other Groups)

Out-Groups	Measures	Settings
Racial and Ethnic Groups	Implicit bias	Countries
Religious Groups	Explicit bias	Communities
Nationalism	Behavioral measures	Cross-National
Linguistic differences	Hostility	Organizations
Tribal differences		Artificially induced group memberships
Political Parties		

# Activity

- Choose one of the following topics:
  - Attitudes toward new technologies
  - Political participation
  - Work engagement
- Identify two different:
  - populations you might study
  - Types of things you might measure
  - Settings where you might find study this

# Types of Data



## Designed Data

- Data processes that are explicitly designed to measure things
- Terminology primarily used in survey research, but concept easily applies to many other types of research
- Most control over populations, variables, measurement, research question
- Power to create data that doesn't exist
- BUT: Often expensive

## Organic Data

- Data which already exists for some other use
- Sometime called “found” data
- “Big Data” is generally organic data
- Administrative records are a type of organic data
- Increasingly common in computerized world
- Large number of observations possible for low cost
- Created for some other purpose
- Population and variables may not always measure what you want
- (What you can get versus what you need)

# The Three R's of Research

- **Realism**
  - To realistically describe or understand
    - Qualitative or in-depth interviews excel at this
- **Randomization**
  - Aim is to isolate variables to determine causal relationships
    - Experiments can be included in surveys
    - Surveys can be included in experiments
- **Representation**
  - Aim is to understand matters pertaining to a larger population
    - Surveys are strong at this

Leslie Kish: *Statistical Design for Research*





# Realism and Richness:

- Thick description or rich understanding
- Types of Approaches:
  - Ethnography
  - In-depth, in-person observation
  - (Participant observation)
  - Qualitative semi-structured research
  - Focus groups
  - Semi-structured interviews

© Chase H. Harrison

# Randomization

- Aim is to isolate the effect of a variable
- The effect of X on Y
- All factors kept same except for treatment (X)
- Subjects (respondents) randomly assigned to treatment ( $X_T$ ) or Control ( $X_C$ )



# Representation:

- Aim is to describe or represent a population
- Surveys are strong at this
- Weaknesses of surveys: standardization of measures reduces richness of measures.



# Types of Statistical Inference in Surveys

## Design Based Inference

- Estimates population characteristic in a population
- Requires (near) full sample frame
- Well documented probability-based sample
- All members of target population have known non-zero probability of being selected
- Data are weighted to adjust for differential selection probabilities

## Model Based Inference

- Assumes that characteristics can be explained by other variables
- Assumes all relevant variables are included in model
- No omitted variable bias
- Typical regression (GLM) model taught in statistics classes

# Pros and Cons

## Design Based Inference

- Makes fewest assumptions
- Requires strictest assumptions about sampling and other errors of representation
- Generally more expensive
- Produces population-level statistics

## Model Based Inference

- Weaker sampling requirements
- May only be generalizable to the population under study
- Requires strict assumptions about model
- Generally less expensive data collection
- Often sufficient for studying relationships

# Population Inference in Surveys



## Finite Population Inference

Generalizing only to target population of study

Strict Generalization only possible with probability survey

Precise point-estimates (simple statistics) possible with few assumptions

General model used in producing statistical point-estimates

## Superpopulation Inference

Assumes that populations vary across people, time, and space

Assumes there is some underlying random process that generates the population

Requires thinking about this process (maybe backward and forward)

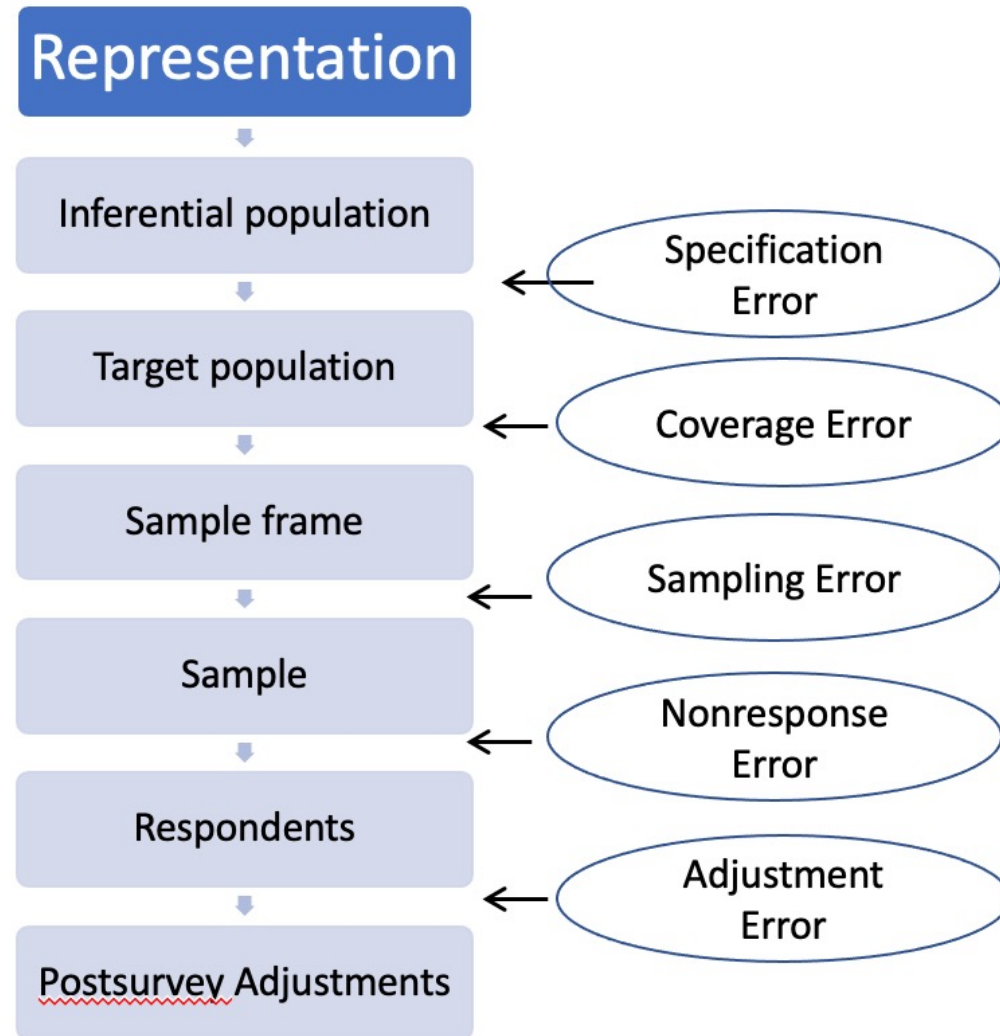
Elucidation may bring model-based assumptions into inference

More typical assumptions of social scientific inquiry

# Finite Populations versus Infinite Superpopulations

---

- Typical model-based inference is to a finite target population
- Target populations often vary across people, space, time, and type
- Try to think about how far beyond your target population you might generalize
- Think about (and explicate) casual processes generating target population (i.e. from superpopulation)



# Types of Surveys

- **Descriptive Surveys**

- Aim is to ***Describe*** something

- **Analytic Surveys**

- Goal is to understand ***Relationships***

# Descriptive Surveys

- *Attempt to make a population estimate or otherwise enumerate something.*
- *Principal goal is to count something*
- *What proportion of a population have a characteristic?*
- *How often do two things occur at the same time?*

## Examples:

- US Census to enumerate the population of the US
- Survey to estimate the proportion of voters who are college students
- Pre-election survey to estimate which voters may be likely to change their minds
- Exit Poll to determine who voted for which candidate and why





# Analytic Surveys

---

- *Attempt to study the relationships and associations between variables*

## **Examples:**

- *Survey to understand relative value of product attributes*
- *Survey to test which likely voter questions do the best job of predicting electoral turnout*

# Types of Surveys

## Cross Sectional

- Survey that measures things at a specific time

## Repeated Cross-Sectional

- Repeated cross sectional surveys asking identical survey questions at different times
- Produces time series data
- Example: Gallup Poll (Trends questions)

## Longitudinal Panel Survey

- Interview sample people multiple times
- Useful for studying change

# Experiments:

- Attempt to study one variable or factor in isolation
- Randomization or control across other factors provides strong evidence for impact of single variable
- Problems with experiments: Difficult to isolate single variable in natural settings.

# Other Types of (Social Science) Experiments

- Randomized Control Trials
- Field experiments
  - RCT's for Program Evaluation
- Natural Experiments
- Conjoint Experiments
  - Often administered in surveys
- Lab experiments

# Survey Experiments

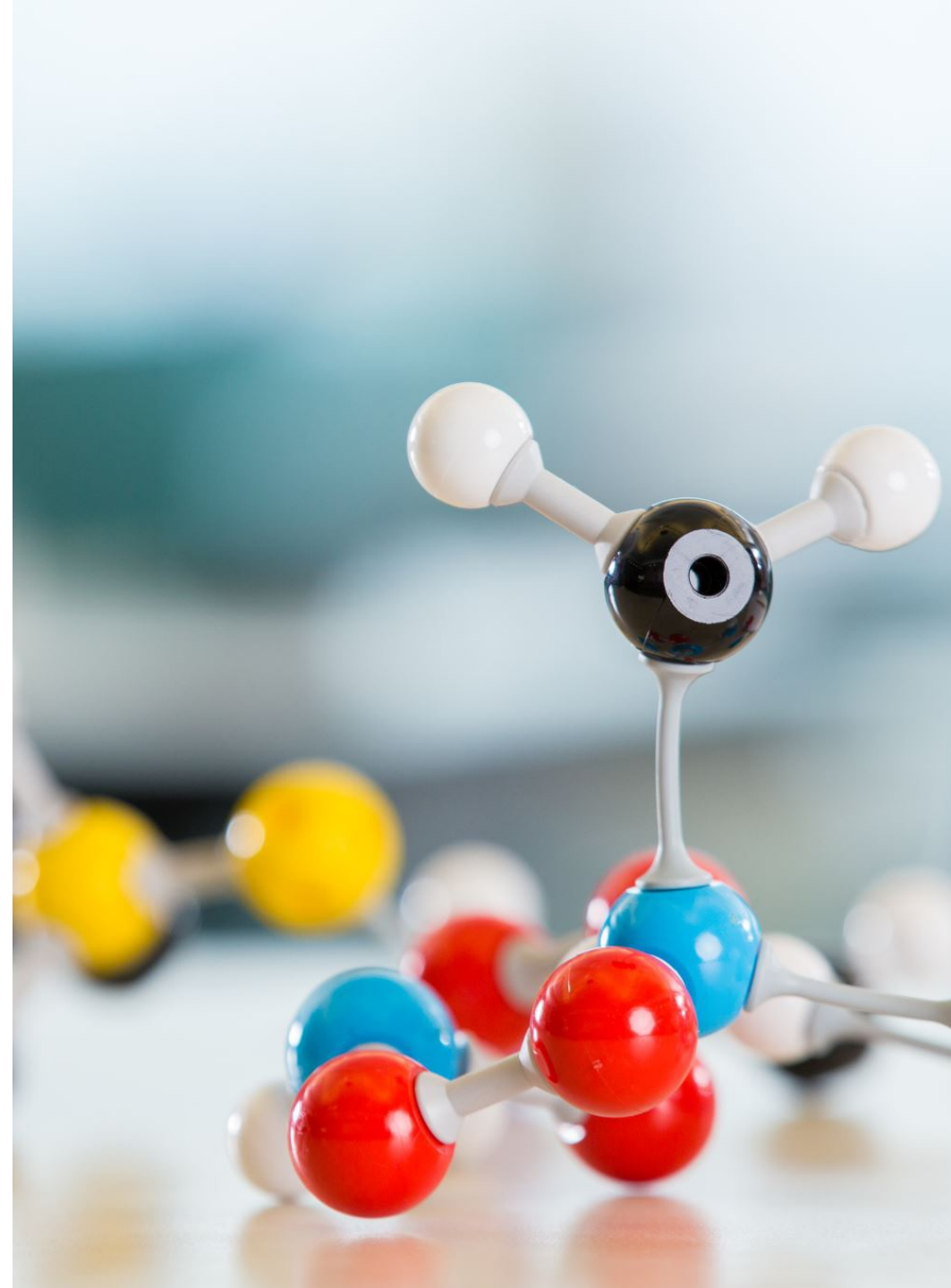
- Designed to provide strong test of a narrow thing (A “treatment”)
  - Respondents are randomly assigned to treatment and control groups
  - Random assignment insures that only difference between treatment and control group is TREATMENT
- Examples:
  - Frames
  - Primes
- Uses:
  - Determine differences in survey designs
  - Message differences
  - Purchase differences (A | B Testing)
  
  - Substantive differences where treatments helps understand causation

# Activity

- Take the same topic you chose in earlier activity.
- Discuss:
  - [REALISM] What you might learn in a realistic research approach such as focus groups or in depth interviews.
  - [RANDOMIZATION] Some measure or variable where you would like to understand the precise influence it has on your topic.
  - [REPRESENTATION] Identify an important statistic you would like to measure and suggest a population you think would be important to measure this in.

# Mixed Methods Research

- Increasingly common approach is both social sciences and applied sciences
- Idea is to use mixture of methods, each designed to focus on its particular strength
- Sometimes use to test extensibility of theories or ideas (triangulation)





# Surveys in Mixed Methods Research

- Qualitative / Historical / Secondary (Quantitative)
- Case Selection
- Focus Groups / Ethnography
- Semi-structured interviews
- Survey Experiments (possibly non-probability)
- Field survey (probability or omnibus items)
- Follow-up interviews
- Quantitative analysis

# Surveys in Mixed Methods Research Example:

Comparative Politics,  
International Relations,  
Development Economics

