



# Introduction to Surveys for Honors Thesis Writers

April 14, 2009

## Overview of Presentation

- › Introduction to Survey Research
- › The Survey Process
- › Survey Resources at Harvard
- › Sampling, Coverage, and Nonresponse
- › Thinking About Modes
- › Question Wording

## The Three R's

### Representation

The research design make inference to a larger population  
 Various population characteristics are represented in research data the same way there are present in population  
**Example:** General population survey to estimate population characteristics

### Realism

A full picture of subjects emerges  
 Relationships between multiple variables and multiple ways of looking at the same variables can be studied  
**Example:** A qualitative case study to evaluate the nature of democracy in a small town with community meetings

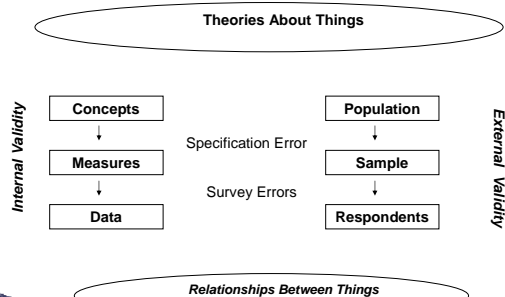
### Randomization

Variables which are not important in model are completely randomized  
 Effects of non-randomized variables can be tested  
**Example:** Randomized clinical trial to test effectiveness of new cancer drug

## Overview of Research Process



### Integrated Example: Surveys and the Research Process



## Two Key Elements of Survey Research

### Good Measures:

Questions or measures impact your ability to study **concepts**  
 Think carefully about the underlying concepts a survey is trying to measure. Do the survey questions do a good job of capturing this?  
**The PSR Tip Sheet on Questionnaire Design** contains good ideas on how to write survey questions.

### Good Samples:

Samples give you the ability to **generalize** your findings  
 Think carefully about the population you are trying to generalize your findings to. Does your sample design do a good job of representing these people?  
**The PSR Tip Sheet on Survey Sampling, Coverage, and Nonresponse** contains thoughts to think about in designing or evaluating a sample.

## Surveys

- ▶ Systematic method of data collection
- ▶ Usually use samples
- ▶ Designed to measure things
  - Attitudes
  - Behaviors
- ▶ Create statistics
  - Descriptive
  - Analytic

## Resources at Harvard

### •Getting Started:

[http://psr.iq.harvard.edu/getting\\_started](http://psr.iq.harvard.edu/getting_started)

### •General Resources:

[http://psr.iq.harvard.edu/internet\\_resources](http://psr.iq.harvard.edu/internet_resources)

### •Tips and Tools:

[http://psr.iq.harvard.edu/type\\_psr\\_resource/psr\\_survey\\_toolbox](http://psr.iq.harvard.edu/type_psr_resource/psr_survey_toolbox)

## Preceptor in Survey Research

*Chase H. Harrison*

*CGIS N-407*

*Office hours: Tuesday 2:30 - 4:30*

*And by Appointment*

[charrison@gov.harvard.edu](mailto:charrison@gov.harvard.edu)

*(617) 384-7251*

## Secondary Analysis of Survey Data

## Secondary Analysis of Survey Data

▶ Question-Level (Micro) Data

▶ Survey-Level Data

## Question Level Data

▶ [Roper Center IPoll](#)

▶ [Polling the Nations](#)

## Survey Level Data

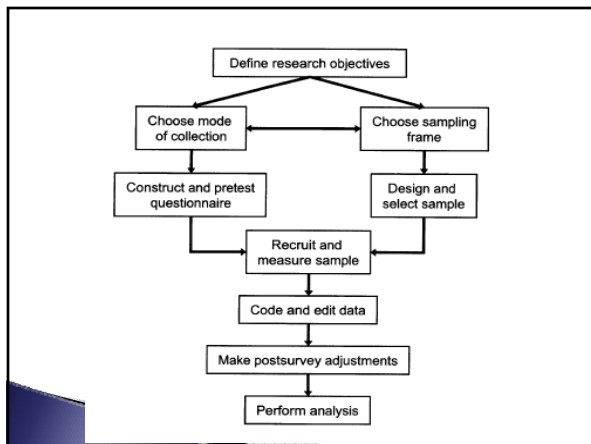
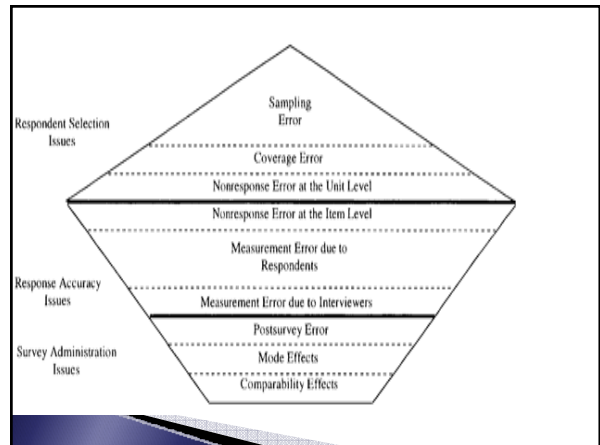
- ▶ [National Election Studies](#)
- ▶ [General Social Survey](#)
- ▶ [Eurobarometer](#)

## Data Archives

- [ICPSR](#)
- [Roper Center](#)
- [Odum Institute](#)
- [European Union Guide to Survey Data Archives and Centers](#)

## Perspectives on Survey Research

- ▶ Survey Design Perspective
  - Steps to use in implementing a survey
- ▶ Survey Quality Perspective
  - Sources of error in surveys
- ▶ Good designs take quality into account
- ▶ Good reports focus on error



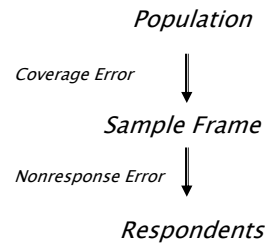
## *Survey Research:*

*Sampling, Coverage and Nonresponse*

## Survey Sampling

- › A **census** attempts to collect data from all members of a population.
- › **Random samples** let you use collect data from a portion of a population and use sampling statistics to generalize your findings to a large population.

## Survey Sampling



## Sample Error

- Based on Statistical Theory
- Describes Variability
- Applies From Respondents to Sample Frame

## Types of Samples

- › **Random (Probability) Samples:**
  - Based on probability theory
  - Allow generalization
  - Sample statistics can be calculated
    - Sample records are drawn from a well-specified frame
    - Sample records are drawn according to random procedures
    - Each sample record has a known probability of selection
- › **Non-Random Samples:**
  - Do not meet the above criteria
  - Inference can only be made to itself
  - Sample error **cannot** be calculated
    - Frame is not well specified
    - Benefits:
      - Convenient and cost effective
      - Can be used for idea generation

## Sample Frames

- ›List or a set of procedures
- ›Records Don't Have to Lead to Respondents
- ›Sometimes Requires Two Stages of Selection

## Coverage Error

- ›People excluded from sample frame
- ›Typical sampling statistics assume no coverage error
- ›Bias:
  - Proportion Excluded
  - Differences Between Excluded and Included

## Nonresponse Error

### •Sample Members Who Do Not Respond

#### •Reasons:

- Unable
- Unavailable
- Unwilling

#### •Bias:

- Proportion Excluded
- Differences Between Excluded and Included

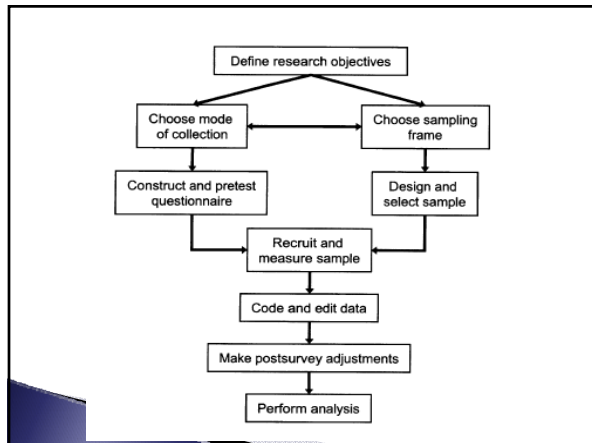
## Addressing Nonresponse

### •Up Front

- Call Backs
- Refusal Conversions
- Incentives

### •Back End

- Weighting
- Statistical Adjustments



## Think About the Mode

### •Self Administered

- Internet
- Mail

### •Interviewer Administered

- Telephone
- In-Person

## Modes

### ▶ Traditional Modes:

- Face-to-face
- Mail
- Telephone

### ▶ Web

- Is it a mode or a method of delivery?
- Sample frame issues confuse mode issue

## Alphabet Soup Modes

### ▶ CAPI Computer Assisted Personal Interviewing

### ▶ CASI Computer Assisted Self Interviewing

- SAQ Self Administered Questionnaire
- CSAQ Computerized Self-Administered Questionnaire

- ACASI Audio Computer Assisted Self Interviewing
- TCASI Text Computer Assisted Self Interviewing
- VCASI Video Computer Assisted Self Interviewing

### ▶ CATI Computer Assisted Telephone Interviewing

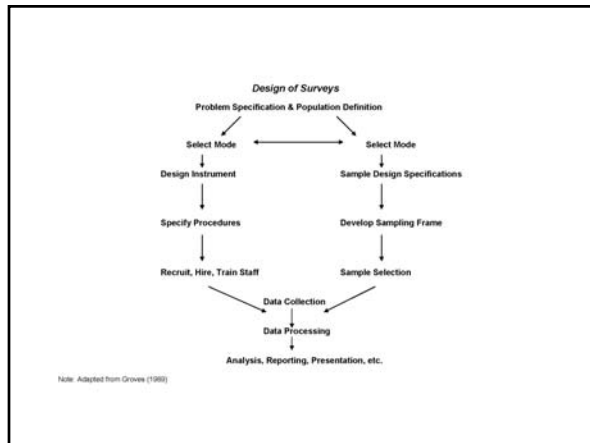
- IVR Interactive Voice Response
- T-ACASI Telephone Audio Computer Assisted Self Administered Interview
- TDE Touch Tone Data Entry

## Where is This Heading?

- ▶ Multiple Modes Within Instrument
  - All respondents receive same questionnaire administered the same way
  - Administration uses different modes:
  - For example:
    - FTF – VCASI – CATI Follow-Up
    - Phone – Mail – Phone
- ▶ Multiple Modes For Different Sample Frames
  - Telephone contact with WWW SAQ
  - Telephone contact with CATI
  - Face-to-Face oversample

## Where is This Heading?

- ▶ Multiple Modes For Different Respondents
  - Contact method varies by respondent preference
  - Contact method varies by difficulty of contact
  - Sub-sampled respondents with expensive modes
- ▶ Multiple Modes For Recruitment or Follow Up
  - Mail pre-notification letters for phone
  - Telephone reminders for mail or web
  - For example:
    - FTF – VCASI – CATI Follow-Up
    - Phone – Mail – Phone



## Sample Considerations

- ▶ Nature of population
  - Are all population members available through mode?
- ▶ Nature of unit
  - Screening for respondents is better with interviewers
- ▶ Nature of record
  - If sample record corresponds to unit, less screening required

## Sample Considerations

- Identification of Valid sample records
- Does this require an interviewer?
  - Can respondents do this themselves through screening?
- ▶ Clustering
- Are records clustered geographically?
  - i.e. face to face interviews may be feasible
  - Otherwise face-to-face reduces sample efficiency

## Measurement Considerations

- ▶ Standardization
  - To what degree does the survey experience need to be the same for all respondents
- ▶ Supervision
  - How much supervision do interviewers or respondents need
- ▶ Protocol Clarification
  - If questions arise during the survey process, how will they be addressed
    - Respondent questions
    - Interviewer questions

## Measurement Considerations

- ▶ Item Nonresponse
  - Higher without interviewers
- ▶ Unit Nonresponse
  - Higher without interviewers
- ▶ Bias with sensitive questions
  - Lower without interviewers

## Measurement Considerations

- ▶ Primacy Effects
  - Respondents answer first thing they see
  - More prevalent in SAQ's
- ▶ Recency Effects
  - Respondents answer the last thing they heard
  - More prevalent in interviewer-administered Q's
- ▶ Randomization or Rotation of Response choices
  - Reduces primacy and recency effects in answer categories
  - Usually requires computer administration

## Measurement Considerations

- ▶ Randomization or Rotation of Response choices
  - Reduces primacy and recency effects in answer categories
  - Rotate if there is an order to the question
  - Randomize if no order

## Measurement Considerations

- ▶ Nature of Information
  - Some information is available from administrative records
  - Sometimes, multiple respondents are needed to answer questionnaires
- ▶ Nature of Respondents
  - Literacy and sophistication have implications for SAQ's
- ▶ Nature of Measures
  - Is audio important?
  - Visual stimulus?
  - Reading?

## *Survey Research: Questionnaires and Questions*

### Surveys and the Research Process



## ***Survey Questions***

- Measure underlying concept
- Don't measure other concepts
- Mean the same thing to all respondents

## ***Don't Reinvent the Wheel***

- Existing Questions Often Preferable
- Trend Data Requires Identical Questions
- Sources of existing Questions
  - [The General Social Survey](#)
  - [The National Election Survey](#)
  - [iPoll \(database of 500,000 polls\)](#)
  - [Polling the Nations:](#)

## ***Pre-Test Your Survey***

- Administer Questionnaire to Small Sample
  - Convenience Sample is Okay
- Have Preceptor Review Your Questionnaire
- Question Understanding Aid
  - <http://mnemosyne.csl.psy.memphis.edu/QUAID/quaidindex.html>

## ***General Considerations***

- Keep the Questionnaire SHORT
- Keep Questions SIMPLE
- Keep Question Order in Mind
  - Earlier Questions Can Influence Later Questions
  - Sensitive Questions Are Best at Back
- Provide Clear Instructions
- Filter and Branch

## ***Types of Questions***

- Open End
- Closed End
  - Rating Scale

## ***Rating Scales***

- Five to Seven Categories
- Provide Middle Category
- Label Scale With Clear Unambiguous Words
- Agree/Disagree Scales Can be Problematic



### ***Things Good Questions Avoid***

- › Technical Terms and Jargon
- › Vague or Imprecise Terms
- › Complex Sentences
- › Double-Barreled Wording
- › Leading or Emotional Language

### ***Rules for Answer Scales***

- › Scale Categories are Ordinal
- › Categories Are Mutually Exclusive
- › Categories Anticipate All Answers

## **Introduction to Surveys for Honors Thesis Writers**

Questions and Discussion