Survey Research
Part 1
Fundamentals of Survey Research
Nicolas Barreto, PhD MPH
Objectives

• Define the goals of a survey
• Basics of survey “error”
• Basics of survey design
Talking about surveys – philosophy of science

- **Objectivism:** Truth exists regardless of our perceptions
- **Constructivism:** Truth is our perceptions, interpretations, and social interactions
Talking about survey methods – A warning

- Multiple levels of discussion
- Words may have different meaning depending on level
  - Ex. reliability
    - Item vs Scale
What you can do with surveys?

• Honestly?
  • Anything that involves measurement that is not “independently confirmable”;
  • i.e., self-report

• Works for qualitative and quantitative research such as:
  • Measurement development
  • Cross-sectional/longitudinal
  • Quasi-experimental
  • Mixed methods
  • Experimental
Why NOT use a survey?

• Short answer?
  • Can you even do a survey “correctly”?  
    • Part art as well as science
  • Asking people questions...
    • People lie, misremember, forget, misread, misunderstand, accidentally lie, not know and answer anyway...etc.
  • Statistical/Methodological issues
    • Various forms of bias to consider
Why use a survey?

• Short answer?
  • *If done correctly* – surveys can provide
    • Accurate descriptions of a population
    • Estimates of parameters and distributions
    • Estimates without the whole population
    • Information quickly and efficiently
    • Generally cheaper
Goals of a survey

• Collect information to answer a question
• Could be a question about *anything*
  • Past/present/future
  • Events/behaviors
  • Self/others
  • People/things/processes
  • Evaluations/attitudes
Two core survey parts

• Measurement
  • What are you measuring?
  • Are you measuring what you think you’re measuring?

• Representation
  • Who are you measuring?
  • Are you measuring who you think you’re measuring?
A survey “done correctly”

• The core parts are also where the errors happen
  • Measurement
    • Validity/reliability, processing
  • Representation
    • Convergence, sampling, nonresponse, adjustment
Basics of Survey
Opportunities for error

Measurement (internal) error

Represented in the diagram:
- Construct
- Measure
- Response
- Edited data

Validity
Measurement
Processing

Coverage
Sampling
Nonresponse
Adjustment

Target population
Sampling frame
Sample
Respondents
Post-survey adjustments

Survey statistic

Representation (external) error
Measurement (internal) error

- Sometimes called “internal validity”
- Difference between
  - What the “true” value is vs the value found in the survey?
  - What was measured vs what was intended?
Measurement (internal) error

- **Validity (Accuracy)** –
  - Whether an assessment measures what it is supposed to measure.

- **Reliability (Precision)** –
  - Measurement yields consistent scores.

- **Processing** —
  - Response recording includes interpreting respondent intent.
Validity vs Reliability

• Validity –
  • A matter of degree
  • An item or scale are not valid, the data collected can be

• Reliability –
  • Item(s) – does the measure come out the “same”?  
  • Scale – multiple items repeatedly assessing the “same” thing
    • Testing scale reliability = Cronbach’s \( \alpha \)
Side thing: Cronbach’s $\alpha$

- Cronbach’s $\alpha$: internal consistency of latent variables
- What is “good”?
  - Rules of thumb are variable
  - Generally? Think the three bears
    - Not too low and not too high
    - $0.70 > \alpha > 0.95$
Types of validity (not all of ‘em...)

• Construct
• Convergent
• Divergent
• Criterion
• Predictive
• Content
• Nomological

• Experimental
• Face
• Incremental
• Structural
• Local
• Generalization
• Postdictive
Causes of measurement (internal) error

- Instrument sensitivity
- Primacy effect
- Recency effect
- Social desirability effect
- Attrition
- Unclear directions
- Ambiguity in questions
- Inadequate time limits
- Inappropriate level of difficulty
- Identifiable pattern of answers
- Administration/scoring errors
Representation (external) error

- Sometimes called “external validity”
- Generalizability
  - Do the results extend to other groups, settings, or conditions beyond those in the study
Representation (external) error

• Coverage (target vs. frame populations) —
  • some within sampling frame ineligible
  • some in the target population will not be captured.

• Sampling —
  • randomness is crucial to sampling.

• Nonresponse —
  • What to do with missingness.

• Adjustments —
  • Methods for generalizability (weight, stratify, etc.)
Gathering respondents

• Primary considerations
  • Theory (goals)
    • Generalization *from* whom
    • Generalization *to* whom
    • Avoid generalization altogether?
  • Practical
    • Money – get the data
    • Time – to do the measuring
    • Access – get the respondents
Causes of representation (external) error

• Non-response bias
• Selection-treatment interaction
• Treatment effects
• Low response rates

• Treatment effects
  • Location
  • Setting
  • Implementation
  • History
  • Placebo
  • Diffusion
  • Hawthorne
Basics of Survey Design
Survey design – Things to think on

• Primary way avoid error
  • Both measurement and representation
• Pros vs cons
  • Really accuracy vs cost
• Three considerations
  • What to measure
  • Who to measure
  • How to measure
What to measure – Vocab

• Item = single question
  • Generally, a single column in raw data set (check all that apply ruins this...)

• Scale = some # of items used to measure a variable
  • Can be combined however you want (sum, mean, other)

• Variable = concept you want to use/measure/observe
What to measure – Items and Scales

• Examples
  • Item
    • Demographics (age, race, year in school etc.)
    • Direct (confirmable) questions (year Dx, year Tx, cost, etc.)
  • Scale
    • Concepts (ratings, evaluations, preferences, attitudes, feelings)
What to measure – Items

• Item/question considerations
  • Question type
    • Likert, semantic differential, uni- vs bipolar, open vs closed
  • Question wording
    • Length, reading level
  • Question answers
    • Response options, objective vs subjective
What to measure – Scales

• Why ask multiple questions about one thing?
  • Think triangulation

• Two options
  • Use extant materials
    • Pro: easier
    • Con: maybe what you want to measure doesn’t exist
  • Create your own
    • Pro: you will measure what you mean to
    • Con: requires lots of steps to be “robust”/“valid”
Who to measure

• Respondents
  • Ideally – the whole population of interest
  • Practically – a sample of the population of interest
  • Efficiently – a sample of the population of interest you can get

• How to pick?
  • Define the ideal, determine the practically available sample, determine what can efficiently be done.
Who to measure – Theory

• Who you are interested in?
• Is there a comparison of interest?
• Variety of information you want?
• Amount of information you want?
Who to measure – Reality

- Economic and logistical concerns
  - Time it takes
  - Money it takes
  - Participation it takes
How to measure

• Which method(s) you use dictate the kinds of data you may get
• Respondents have resources that you are using
How to measure – Gathering responses

• Myriad of options with different pros and cons
  • In-person
    • Most expensive, but ensure quality of data (less missing)
  • Phone
    • Moderately expensive and some may not have phones (also which #?) but ensure data quality
  • Mail
    • Relatively cheap, but there are response rate and data quality issues
  • Web
    • Cheapest and most flexible but response rate and data quality issues
How to measure – “Tailored” designs

• Mixed mode data collection
  • Some combo of multiple methods
  • Covers weakness of each other
  • Limits the variety that some of them provide
    • Ex. Thermometer questions
How to measure – respondent considerations

• All methods formats have respondent considerations
  • Comprehension and retrieval
  • Time burden
  • Cognitive burden
• Respondents have a “health bar”
How to measure – survey set up

• Order matters!
  • Health bar (some questions are easier than others)
  • Participants pay attention to questions

• Question format
  • The font, consistency, and tidiness of the survey
Recommended reading?

Questions?