SOC497/L: SOCIOLOGY RESEARCH METHODS

Operationalization:

Making Accurate Measures

Ellis Godard

Guy who invented the clock: there will be 12 numbers on it

Friend: so the day will be divided into 12 segments?

Inventor: no, 24

Friend: so will the day start at 1

Inventor: the day will start at the 12, which is at night

Friend:

Inventor: the 6 means 30

2

Admin Operations Prcsn Reliability Validity Threats Etc

Outline for Today

- **Definitions**
- **Operations:** Meanings & Choices
- **Accuracy:** Precision, Reliability, Validity
  - Tradeoffs & Lab: Internal Threats
- Next time: Levels of Measurement (LOM)
  - Choices & Implications

That's why I have trust issues !

Admin Operations Prcsn Reliability Validity Threats Etc

Which love did/will you count?

A. Lustful 5%

B. Romantic 75%

C. Platonic (friendly) 10%

D. Familial 10%

E. CSUN (school spirit) 0%

F. Other 0%

30

3 Kinds of Definitions

- Real: … Think about the "love" lab…
  - **Avoid** quagmire of "real" meaning; Make a choice & move on!
- Nominal: represents consensus/convention about term
  - Soc dictionaries, everyday understanding, scholarly literature
  - No claim that it represents "real" entity's essence
  - Don't confuse w/ nominal measurement
- Operational: specify precisely how concept *measured*
  - What "counts"? E.g. SES = Low, middle, high? based on income?
  - That's what matters for research, regardless of reality or convention
- All share common purpose
  - Permit comparison of observations
  - Absolutes or constants don't help; must *vary*

Q1. Which kind of definition should you avoid constructing?

A. Real 10%

B. Nominal 2%

C. Operational 0%

D. All of the above 10%

E. None of the above 79%

Choices in Operationalization

- Choices to Make
  - Level of Measurement
  - Range of values & Extremes reached
  - Stability of measures
    - Some change during collection, entry, coding, analysis…
- Guidelines for choices
  - Pragmatic goals: address given task, target known audience
  - Practical: given resources (data, staff, budget, equipment, time)
  - Scientific: maximize information, reduce biases/risks

Next

"Accurate"

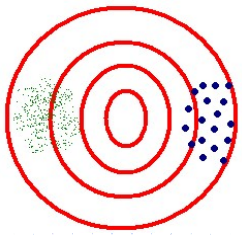
SOC497 ☺

Accuracy

- Whether get the "true" measure
  - but not oft used b/c vague (~ "real")
- Has 3 parts; 1st = precision
  - Accuracy = closeness to "the truth" (?)
  - Precision = finiteness of distinctions
    - How specific are the measurement categories?
    - agree/disagree versus strong/agree/neut/disagree/strongdis

Precision

- 2 sets of dots, about equal in two ways:
  - How close they are to center
  - How distributed they are from each other
- But left more precise
  - Those on the right are more blunt or broad



Less precision may be…

- Sufficient for procedures
  - Crosstabs: "grew up in 60s" may be enough
- Necessary given resources
  - If "in her 40s" is enough, additional effort = waste
- More accurate
  - "born in Stowe, VT" more precise than "born in new England" unless born in Boston, MA
  - Precision risks errors

Reliable Measurements…

- Produces stable & consistent results
  - Reliability: extent to which a procedure or technique, if repeatedly applied to same observations, would yield same results each time
- Do not ensure accuracy more than precision
  - Consistent inaccuracy = bias
    - left of target
    - underreport income
    - overreport # partners
    - bathroom scale 5 pounds low
    - watch 5 min off
  - Misinterpretation always possible
    - Consistent misinterpretation is reliable (just not valid)

Reliable Exams

- Exam Formats
  - Essays? multiple-choice! true-false!
- Asking questions reduces reliability
  - Esp. if separate sections
  - Ambiguities/difficulties should be encountered equally by everyone, and resolved uniformly later
  - Otherwise, reliability of exam in doubt.

Threats to Reliability

- Beyond area of knowledge
  - satisfied w/ China's current policy towards Albania?
- Can't recall
  - how many times ever been to church?
- Question complicated or unclear
- IVR/R interaction problematic
  - 2 IVRs ask for R's sexual behavior -> reliable
  - M&F IVRs ask R for # of sexual partners -> unreliable
- Only one observer/researcher
  - b/c no guard against subjectivity

7 Ways to Check Reliability

In text… .

1. Test-Retest (aka "stability reliability")
   - Reliable across time?
2. Split-Half (aka "representative reliability")
   - Same answer when applied to subgroups?
   - Split sample in half, give each same questions
   - Reduce an index to two parts, & expect parts to correlate
3. Use Established Measures (aka "equivalence reliability")
   - Several indicators consistent
   - Use what has worked in the past – SAT, GRE

7 Ways to Check Reliability, cont'd

Others… .

4. Replication
5. Clarity and specificity in "instrument/tool"
   - also, given case/respondent/record, IV/Iver
6. Training (& practice) – this class
7. Statistical: Cronbach's alpha (esp. for Indices!)
   - Roughly, the percent of times the variables "agree" w/ joining
   - Similar to a correlation, though *not* same formula as Pearson's r

Q2. Which relies on conventions?

A. Chronbach's alpha
5%

B. Clarity in instrument
0%

C. Established Measures
95%

D. Split-Half
0%

E. Test-Retest
0%

6 Types of Validity

- Q: Are we measuring what we think we are?

1. Face Validity
   - A measurement we agree is valid on its face
   - Seems to relate to concept, common agreements
2. Formal Agreements
   - E.g. Census defines family, GSS defines rincome
3. Criterion validity
   - Also called predictive validity
   - External criterion = behavioral confirmation in *past*
   - E.g. accidents & driver tests, college grades & SAT

Q3. Driving tests have…

A. Construct validity
   5%

B. Content validity
   10%

C. Continuing Construct validity
   0%

D. Criterion validity
   85%

E. None of the above
   0%

6 Types, cont'd

4. Construct validity
   - Based on a logical relationship among variables
   - Confirm w/ other measure *in data*
   - Validity based on weight of evidence
   - E.g. marital satisfaction to check whether R cheats
5. Continuing construct validity
   - External but ongoing – like 3+4
   - Satisfied customers at a particular car dealership
     - Expect to recommend the dealership to others
     - Source and consequence of them coming

External vs. Internal Validity

- Internal Validity
  - accuracy of conclusions drawn from a study
  - Explore 15 threats* next…
- External validity
  - Concerns the generalizability of findings
  - w/ sampling, after midterm

Frey, Botan, and Kreps, *Investigating Communication*, Allyn & Bacon, 2000, pp. 119-125

6 Types, cont'd

6. Content validity
   - Whether a measure covers the (full) range of meanings included within a concept
     - Tests of math ability can't just look at addition
     - "Prejudice" has problems if not clear what kind
     - "Love" not clear if not clear what kind
     - Feminism – broad idea about commitment to full equality in work, family, arts, politics, and more
       - "Should men & women get equal pay for equal work?"
       - "Should mean and women share household tasks?"

Threats to Internal Validity

- 4 General categories (15 specific threats)
  - How Research is Conducted (x3)
  - When Research is Conducted (x4)
  - Researcher Issues (x3)
  - Participant Issues (x5)

How Research is Conducted

- Procedural validity
  - Whether the measurement process is "clean"
  - Accurate and consistent techniques
- Treatment validity
  - Whether the IV is the same as intended
  - To study effects of scary film, must *be scary*
- Environmental validity
  - Cafeteria vs. library
  - Problem in field research (home vs. office)

When Research is Conducted

- History
  - Changes in environment external to study (e.g. news)
- Sleeper effect
  - Effect becomes apparent over the course of time
    - E.g. Reactions to media may not measurable immediately
- Sensitization
  - Tendency for 1 measurement to influence a later 1
  - Respondent becomes aware of *process*
    - Four-part Solomon experiment, w/ no pretests for 2
- Data Analysis
  - Improper procedures to analyze data
    - Remember 364? ☺

Researcher Issues

- Personal Attribute Effect
  - Interviewer characteristics interfere w/ measurement
- Unintentional Expectancy
  - Accidentally conveys what you want to see
    - E.g. smile in person, laugh / raise voice on phone
- Observational Bias
  - Observer drift – become inconsistent
  - Observer bias – from knowledge of research

Participant Issues

- Hawthorne Effect
  - Change in behavior due to being observed
  - Observing the observer – aware of intent
- Selection Bias
  - Comparison assumes respondents are who we say
    - Self-selection bias (groups formed by relevant var.)
    - Regression to the mean – less extreme in later measures
    - Ceiling effect – Rs already high; can't assess treatment effect
- Participant Mortality / Attrition
  - Loss of Rs during study – death, departure, disinterest
- Participant Maturation
  - Internal changes *within* people during a study
    - E.g. competence increaes, vision worsens
- Inter-Participant Bias (or inter-subject bias)
  - Inc. diffusion of treatment

Q4. What kind of threat to validity is an earthquake?

A. Environmental validity
49%

B. History effect
46%

C. Personal Attribute Effect
2%

D. Selection bias
0%

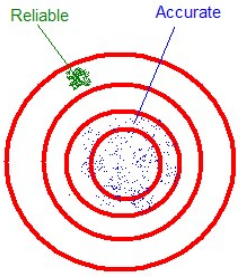
E. Sensitization
2%

Reliability vs Validity

- Consider the measurement "How much do you drink?", retested over time
  - Might get "not at all" at Time 1 & Time 2
  - But what if Time 1 is before 18?
  - What if Time 2 is in very old age?
- The measurement may be reliable, but it's subject to maturation, a threat to internal validity

Reliability vs. Validity *(Illustrated)*

- Consistency vs. Closeness



36

Q5. Which is *not* part of accuracy that we measure and protect?

A. Precision
0%

B. Reliability
9%

C. Validity
0%

D. Truth
91%

E. All of the above
0%

SOC497 @ CSUN w/ Ellis Godard

Reliability vs. Validity Tradeoff

- Often sacrifice one for the other
  - if specific enough to be consistent, may zero in on something not of interest
- Don't want to generalize a false conclusion
  - Research only valid for a given sample, context, or time is of limited use

SOC497 @ CSUN w/ Ellis Godard 37

Team Scores

Points	Team	Points	Team

SOC497 @ CSUN w/ Ellis Godard 40

Today's Lab Exercise

- Handout on the web
  - Threats to internal validity
- Apply lecture to research problems
  - Some real, some hypothetical
- For each scenario, unless otherwise specified, **give the general category *and* specific type** of threat implied

SOC497 @ CSUN w/ Ellis Godard 38