Lecture 6

A. Measurement

Definition: Assigning numbers to concepts for measurement purpose or test hypothesis. Translate theory to real world 'application'

Measurement: foundation of good research

i) Characteristics
   a. Validity
   b. Reliability

ii) Steps
   a. Identification of concept. Eg creativity, satisfaction
   b. Develop construct
      i. Measuring one concept by asking various questions (variables)
      ii. Regards to abstract concepts
      iii. Can't just ask how loyal are you as a customer? It is subjective.
   c. Conceptual and operational definition
      i. Review literature for conceptual definition – what is the theoretical definition
      ii. Should match with measurement items, testing criteria
   d. Develop measurement scale
      i. Choose existing scale (likert), (semantic diffraction-opposite adjectives)
      ii. Assignment of numbers to the measurement of a concept-quantitative data

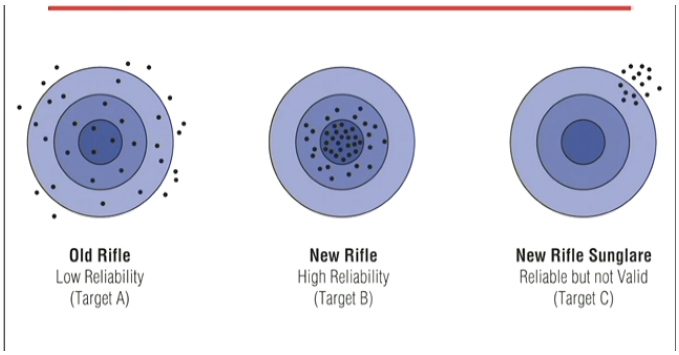
*** Categories of scale: Non metric (Nominal & ordinal); Metric (interval & ratio)

Nominal	Ordinal
Useful for identification and classification	Identification and classification
Relative position can't be indicated (no order)	Relative position
Magnitude of differences cant be compared	Magnitude of differences cant be compared (by how much more like it)
Ratio of scale values cant be compared. Zero doesn't have meaning	Ratios of scale values cant be compared
Country of origin, Gender	Level of education, brand preference

Interval	Ratio
Identification and classification	All apply
Relative position	
Magnitude of differences can be measured. Intervals must be equal	
Ratios of scale values cant be compared	Absolute zero point (cannot have zero)
GPA , Likert	Height, age

*** Average scores for items that measure one construct. We want the scores to be similar, ie give same information

e. Reliability, validity and sensitivity
   i. Validity: measures what we want it to measure. EG Performance – use salary not a good measure as highest pay doesn't always mean highly performing
      1. Face: simplest form, does it seem to be logic/accurate? EG how many children you have? (face validity is enough). Ask anyone
      2. Content: How well measure captures all facets of a construct in the test. EG Performance, using only grades isn't enough. Refer to literature/ expert judgement to check content validity
      3. Convergent: Extent to which measure correlates positively with other measures of the same construct, ie want scores to go in same direction
      4. Discriminant: Measurements should be distinct from each other (regression/correlation)
      5. Criterion: how well does it predict the outcome. SAT scores will determine performance at uni (future outcome-predictive validity) VS concurrent validity – how well does measure support previously validated measurement for the same construct
      6. Construct: extent to which a test measures what it is intended to and allows inferences to be made back to population
   ii. Reliability: consistent results. If test taken number of times, results should be same each time. Contributes to validity.
   → Use internal consistency (measure homogeneity, indicators converge same meaning using split half method=compare 1 half set to other half and coefficient alpha=how close related a set of items are as group >0.7 but don't want it to be to high{high multicolinearity} cause or else the measures become redundant, ie doesn't add anything) & Test-Retest reliability= same scale or measure to the same respondents at 2 points in time to test stability



Old riffle: not reliable and not valid
New riffle: Reliable and valid
New riffle sunglare: Reliable (same area targeted) but not valid (missing central point)

   iii. Sensitivity: More response categories and more items increase sensitivity. EG sensitivity in elevator to not hurt anyone when the doors close or too sensitive that it

opens all the time. EG security in airports- too sensitive to metal?

Measurement error

- Respondent error:
    - Reluctant to express strong feeling (don't use whole scale)
    - Little knowledge (use guess)= no accuracy
    - Fatigue, boredom, anxiety (from long surveys)
- Situational factors:
    - Shouldn't strain
    - If another person present, person can distort responses by joining in or distracting
- Interviewer error
    - Hard to stay neutral or leading
    - Act differently if you like/dislike interviewee
- Instrument error
    - Question confusing, long, hard to understand
    - Poor sampling- non representative of the wider population

B. Measurement analysis

- Multiple Qs to measure one concept
- Data transformation: process of changing original numerical representation of a quantitative value to another value
    - Reverse coding
    - Aggregating data: summation and averaging (need to have convergent validity- scores should be similar, sensitive to extreme scores= mean wont be representative)
- Reliability testing: Cronbach's alpha (>0.7). Closer to 1 higher consistency/correlation. To have 1 all scores would need to be exactly the same
- Validity testing: factor analysis