Chapter 5: Identifying a Good Measurement

Measuring Variables

- Conceptualize: defining different abstract words, such as happiness, depression, attraction, etc., into a precise conceptual definition
    - Conceptual variable: researcher's definition of the variable in question at an abstract level
- Operationalize: taking the conceptual definition and making it so it can be measured and manipulated as a variable in a study
    - Operational variable: represents a researcher's specific decision about how to measure or manipulate the conceptual variable
- **Construct validity**: a measure of how well a variable was measured or manipulated in a study

Three types of operationalization (or measures):

1. Self-report measure: measures a variable by recording peoples answers to verbal questions about themselves in a questionnaire or interview.
    - Sometimes it is more accurate to have a self-report measured filled out by someone who observes behavior versus someone who engages in the behavior
2. Observational measure: measures a variable by recording observable behaviors or physical traces of behaviors.
    - Number of smiles
    - Number of smiley faces sent in text messages (physical traces)
3. Psychological measures: measures a variable by recording biological data such as brain activity, hormone levels, or heart rate.
    - Some people believe physiological measures are best, but think about potential confounds

Operational definitions can be:

1. Categorical: categories (gender, species, grade)
    - Sometimes called nominal: assigning the number 1 to males and 2 to females
        - This is not really measurement, it is merely labeling for purposes of differentiating groups.
2. Quantitative: coded with meaningful numbers (height, weight, temperature)
    a. Ordinal scale: applies when the numerals of a quantitative variable represents a rank order (first place in a race)
        - The distance between the numbers may or may not reflect equal intervals. The first place team could be WAY out in front or barely winning
    b. Interval scale: applies to the numerals of a quantitative variable meets two conditions: 1. represents equal intervals (distances) between levels and 2. There is no "true zero" (a score of zero does not mean "nothing")
        - Examples: IQ test, temperature in Celsius, most questionnaire scales
    c. Ratio scale: applies when the numerals of a quantitative variable have equal intervals and when the value of zero truly means "nothing"

- Weight, income, height

Reliability of Measurement: how consistent a measure is- is your measure consistent over time?

Three types of reliability:

1. Test- retest reliability: the researcher gets consistent results every time he or she uses the measure.
    - Scores should be consistent
    - Can be relevant no matter whether the operationalization is self-report, observational, or psychological
    - Primarily relevant when researchers are measuring construct
2. Interrater reliability: consistent results are obtained no matter who measures or observes
    - Most relevant for observational measures
3. Internal reliability: a study participant gives a consistent pattern of answers, no matter how the researcher has phrased the question

Correlation Coefficient "r": a single number to indicate how close the dots on a scatterplot are to a line drawn through them

- Describes the direction and strength of the association between two variables
- Strong (closer to 1 or -1) when the dots are close to the line and is weak (closer to 0) when the dots are spread out
- Positive: move together
- Negative: move opposite

Cronbach's alpha: correlation based measure, provides information about internal reliability

- The closer the Cronbach's alpha to 1, the better the scales reliability
- Cutoff of 0.7 necessary for measure to be considered reliable
- If internal reliability is lower, the researchers are not justified in combining all the items into one scale

Validity of Measurement

- Face validity: looks as if it should be a good measure
    - Does it appear to measure what it is supposed to measure?
- Content validity: must measure the entire range of a construct
    - Usually evaluated and scored by experts in the content area.
    - a CVI (context validity index) of .80 or more is desirable
    - If we are looking at intelligence, measure to evaluate ability to:
        - Plan
        - Reason

- Think abstractly
- Comprehend complex ideas

- **Criterion validity**: the degree to which content on a test correlates with performance on relevant criterion measures
  - If they do correlate highly, it means that the test has criterion validity
  - Two types of criterion validity:
    - **Concurrent Criterion Validity** = how well performance on a test estimates current performance on some valued measure (criterion)?
      - (e.g. *test of dictionary skills* can *estimate* students' current skills in the actual use of dictionary – observation)
    - **Predictive Criterion Validity** = how well performance on a test predicts future performance on some valued measure (criterion)?
      - (e.g. *reading readiness test* might be used to *predict* students' achievement in reading at the end of a semester
    - Both are only possible **IF** the *predictors* are **VALID**
- Convergent validity: the degree to which a measurement correlates with other established measurements of a construct
  - Correlate less strongly with measures of other distinct constructs
  - A new measure of intelligence should correlate highly with the current standard (the WAIS)
- Divergent validity: the degree to which a measure does not correlate with a measure of a different variable
  - A measure of depression should NOT correlate highly with a measure that has been connected to self-esteem.
  - If it does we have no idea if the new measure is really measuring depression, self-esteem, or a combination of both

Relationship between reliability and validity:

- The validity of a measure is not the same as reliability
- May be less valid than it is reliable, but cannot be more valid than it is reliable
- Reliability does not equal validity

Chapter 6: Describing What People Do: Surveys, and Sampling

Survey: when people are asked about a consumer product

- Survey question formats:
  - Open-ended questions: allows respondents to answer in an any way they see fit
    - Example: Describe your thoughts on the rising cost of parking at OSU, which is scheduled to increase 5% next year.
  - Forced choice format: which people give their opinion by picking the best of two or more options
    - Example: Choose one of the following statements that best connects with your true feelings: