10-29 and 10-31 notes (and also probably a bit of 11-5, too)

S and B data (from Chap. 2)

*** No need to read chap. 2, but can use as a backup***

1) Self-Judgment or **Self-Report** (S) Data
    - Definition: a person's evaluation of his or her own personality (giving own opinion of yourself)
    - Usually with questionnaires or surveys, MC, T/F
    - Simple, easy--most frequent way to collect data from people
2) Behavioral (B) Data
    - Definition: information that is carefully and systematically recorded from direct observation and measurement of behavior
    - "most visible indication of an individual's personality is what she does"

Natural B Data vs. Lab B Data

Natural B Data: Based on real life

- Diary and experience-sampling methods (compromise to following people around all day)
- Reports by acquaintances
    - People who know the person report, common with parents/chidlren
- Naturalistic observation
    - Watch in natural environment
- Laboratory B Data: based on Experiments

  Make a situation happen and record behavior (watch what people do)
  – Frustrating situation
  – Attention (continuous performance test for ADHD)

  Examine reactions to subtle aspects of situations (more about this later)

The Nature of Personality Assessment

- More than just measuring traits
- Personality: characteristic patterns of behavior, thought, or emotional experience that exhibit relative consistency across time and situations
    - How do you usually think, feel behave?

--Also measure motives, intentions, goals, strategies, and how people perceive and construct the world

Personality Assessment

- Assessment (from Oltmanns and Emery, 2012, Abnormal Psychology): "gathering and organizing information about a person's behavior"
    - we try to assess personality, etc., behavior only accurate
    - assess people around us all the time informally
- Includes: mood and personality surveys
    - IQ tests
    - Questionnaires asking you about you
    - Behavioral measures
    - Psychoeducational testing (learning disorder)

Personality Tests

- Used by psychologists (experimental and clinical), corporations, and the military

- Omnibus inventories: measure a wide range of traits—anxiousness, depression, masculinity, lying, suicidal
- Some tests are designed to measure one trait (much shorter)
- Most tests provide S data
- Some tests provide B data
    - Minnesota Multiphasic Personality Inventory (MMPI)—also provides S data (along with B)!
    - Implicit Association Test (IAT)—remember bias tests? IE- behavior data about race
    - Intelligence (IQ)
        - Uncomfortable because need to know where one fails to establish limit

Why Assess?

- To precisely get and communicate information about mood, personality and behavior
- To help predict future behavior
    - IE) high on this measure= behave this way in this particular situation?
- To differentiate between normal and abnormal behavior
    - IE) most people wouldn't do this; would be happy about this happy with themselves…
- Can help establish diagnosis, which helps determine course of treatment

Note: In order to have some confidence in testing results administration must be standardized—same testing conditions, same instructions, same time limits, etc.

- E.g long instructions on the SAT

- Reliability and Validity

Test must be both reliable and valid if results are to be accepted

Reliability=consistency of an assessment device

* IQ more stable than depression

Several ways to measure reliability:

- Test-retest method: give a subject the same test twice over a period of time to see if the scores are highly correlated (correspond closely). Problems? **A=A?**
    - still around the same percentile the second time you take the test for neuroticism?
    - Factors could change, people could know what to expect and answer differently
- Equivalent forms method: give the subject two forms of the same test, and see if the scores correlate. Problems? **A=B?**
    - Time consuming for subject, push patience, can be expensive
- Split-halves method: give the subject a test only once, but compare one half of the test to the other. Problems?
    - Under assumption that all are measuring same thing, usually split odd v even

- Using more items gives more accurate/stable results
- *A/2=A/2??*

- Interrater reliability: two raters/examiners score the same test—do their scores agree? Problems?
  - you're A scoring= my A scoring?
  - Used for subjective judgment scoring test (IE SAT writing section)
  - Different bias= different scoring= different results

Measuring Reliability

- Measures of reliability yield a **reliability coefficient**
  --number that expresses how consistent a measure is
  --all reliability coefficients are forms of **correlation coefficient**—number that expresses the strength of the relationship between two variables.
    - Kind of like, how often, how closely are these related? 80% 20%
    - Do tall people always weight more than short people?

Validity

Validity=test's ability to measure what you mean it to measure.
Just a few of the types of validity:
Face validity: the degree to which the test appears to be measuring what it says it's measuring

- appearing to measure something doesn't mean you're actually measuring something
- e.g- a measure of depression contains a true/false question such as "I feel sad"
- items to measure happiness? Conscientiousness?—do I get my homework done on time?
- "I get excited when I meet new people in an unusual situation"

Content validity: how well *individual items*, or questions, measure what they are supposed to measure.

- "I feel hopeless about the future"—measure of suicidality?

Concurrent Validity: how well a measure compares to a (valid) one that already exists.

- Aggression measure vs. obs. rating

Construct validity: test's overall ability to measure a construct—(hypothetical or theoretical way to explain something, such as a trait, emotion, behavior, or motive).

- Can I validly call my measure as a measure of a particular construct (am I operationalizing it properly)?
- Angry- red face, physically aroused
- Test it could be FACE valid but not CONSTRUCT valid—appears to be measuring something, but not actually measuring what we think it is

Predictive validity: how well a test predicts future (or past) behavior.

- Test of impulsivity used to predict whether employees at a railroad would be injured in the next 12 months