Psychological Testing

Measurement

- A psychological test is a kind of measuring instrument, like a ruler or scale.
- As such, it must meet the requirements of any good instrument:
  - Validity
  - Reliability
  - Accuracy

Validity

- Validity is the extent to which a test measures what it is intended to measure.
- Many ways of assessing validity exist. Some of them are:
  - Face validity
  - Concurrent validity
  - Predictive validity
Face Validity

- A measure has face validity if the items on the test appear (on the “face of it”) to measure what the test is supposed to measure.
- A weak form of validity. Some tests seem to have little face validity and yet are valid measures by other criteria.

Concurrent Validity

- A test has concurrent validity if its results agree with those of an established test (one whose validity is accepted).
- Procedure
  - Administer the new test.
  - Administer the established test.
  - Compute Pearson $r$
- A high Pearson $r$ indicates that the test is valid.

Predictive Validity

- A test has predictive validity if its results make predictions that agree with observation.
- Procedure
  - Administer test to a group of individuals.
  - Collect data on performance related to variable measured by the test.
  - Compute Pearson $r$.
- A high Pearson $r$ indicates good predictive validity.
Reliability

• An instrument is reliable to the extent that repeated measurements of the same thing under the same conditions agree.
• Two kinds of reliability can be established.
  – Test-retest reliability
  – Split-half reliability

Test-Retest Reliability

• Procedure
  – Administer the test
  – Wait awhile, then administer it again to the same individuals.
  – Compute the correlation (Pearson r) between the two results.
• Test is considered reliable if the correlation is high (e.g., above +.90)

Problems With Test-Retest Reliability

• May overestimate a test’s reliability if people remember how they responded to items on first administration and simply answer that way again. This can be combated by using an alternate form on second administration.
• May underestimate a test’s reliability if the characteristic being measured changes substantially between administrations of the test.
Split-Half Reliability

• Procedure
  - Administer the test once.
  - Randomly split items into two groups with half the items in each group (split halves).
  - Score the split halves separately.
  - Compute Pearson $r$ on the resulting pairs of scores.
• Reliability is high to the extent that Pearson $r$ is high (e.g., above .90).

Advantages/Disadvantage of Split-Half Reliability

• Advantages
  - Requires only one testing session.
  - Eliminates the possibility that the variable being measured will change between measurements.
• Disadvantage
  - No guarantee that the two “split halves” are equivalent. If not, then method underestimates the reliability of the test.

Accuracy

• A measure is accurate to the extent that its results agree with a known standard.
• However, in psychological tests there are no “known” standards available to compare results to.
• Instead, results are standardized.
• In effect, standardization uses group data to create a kind of standard based on average group performance.
Standardization

- Procedure
  - Administer the test to a large standardization sample chosen to be representative of the target population.
  - Score the test to get the raw scores.
  - Compute the mean (average) and standard deviation of the scores.
  - Use these to create standard scores having a specified mean and standard deviation.
- Results of tests measuring the same variable and standardized to the same values can be compared.

Examples of Psychological Tests

- Intelligence Tests
- Personality Tests
- Aptitude Tests
- Interest Inventory

Intelligence Tests

- Several standardized intelligence tests are available.
- All are designed to measure the kinds of intelligence that are required for good academic performance.
- We will be examining these in more detail in another lecture.
Personality Tests

- These measure various aspects of personality.
- Objective Tests
  - Can be objectively scored.
  - Example: MMPI
- Projective Tests
  - Allow person to “project” something of themselves into their responses.
  - Scoring is often subjective.
  - Examples: Rorschach Inkblot Test, Thematic Apperception Test

Aptitude Tests

- Aptitude tests attempt to predict whether someone will be able to acquire some new skill or ability with relative ease or difficulty.
- All aptitude tests actually measure a person’s current abilities that are foundational to the skill or ability to be acquired. That is, all aptitude tests are achievement tests.

Interest Inventory

- I like to mention this one as you may find it useful to help decide on a career.
- Test compares your interests to those of people who are happy and successful in their fields of work.
- If your interests match those of people in a given field, then you might want to look into that field.
- Caution! Does not determine your aptitude for that sort of work!