Measuring Instruments

- How data collection happens
- Understanding Relationships among:
  - Constructs
  - Variables
  - Instruments

Constructs

- Cannot be observed directly
- Are invented (inferred) to explain behaviors
  - Intelligence
  - Teacher effectiveness
  - Creativity
  - Motivation
- Measure by identifying scores or values constructs can assume

Variables

- Must have two values or scores
- Can be categorical or quantitative
- Can be dependent or independent

Measuring Scales

- Nominal
- Ordinal
- Interval
- Ratio

Nominal Variables

- Categorical variables
- Lowest level of measurement
- Members of category must have at least one common characteristic

Ordinal Variable

- Not only classify but also rank from higest to lowest
- Show comparisons between and among items
- Do not show how much difference (intervals are not equal)

Interval Variables

- Have same characteristics of nominal and ordinal variables
- Have equal intervals
- Score of zero does not indicate absence of knowledge
- Score of 100 does not indicate total knowledge

Ratio Variables

- Highest, most precise level of measure
- Characteristics of others
- Has true zero point
- Encompasses mainly physical measures
  - Time
  - Weight
  - Distance
  - Speed

Combinations of Variables in Research

- Two or more quantitative
- One categorical – one quantitative
- Two or more categorical

Characteristics of Measures Collecting Data

- Use an existing instrument
- Construct and original instrument
- Record natural events or use existing data

Which Measure ?

- Consider
  - The variables in study
  - Alternative instruments
  - Time

Using Standardize Instrument

- Saves time
- Developed by experts
- Results from studies using same instrument can be compared
- Thousands of published instruments are available

Definition of Terms

- Test –subset of assessment, a formal, systematic procedure for gathering information about cognitive and affective characteristics
- Standardized test – administered, scored and interpreted the same every time
- Assessment – broader term than test – collecting, synthesizing, and interpretation formal or informal data
- Measurement – the process of quantifying after data collection

Data Collection

- Paper and pencil
- Observations
- Interviews

Paper and Pencil

- Selection methods – predominate type used in qualitative research
  - True / False
  - Multiple choice
  - Matching
  - Supply methods
  - Fill in blank
  - Short answer
  - Essay

Interpreting the Instruments

- Raw scores
  - Number of items person scored
  - Basic data analyzed
- Norm - Referenced
  - Compares one person to all others taking instruments
  - Often reported as percentile rank or stanines

Interpreting the Instruments

- Criterions – Referenced
  - Compares scores against predetermined levels of performance
- Self Referenced
  - Compares a students performance over time to determine improvement

Types of Measuring Instruments

- Cognitive – achievement
- Aptitude – prediction of performance
- Affective – feelings, values, attitudes
- Attitude Scale – determine what person believes, perceives or feels
- Interest Inventories – likes on dislikes
- Value Tests – old – measures values
- Personality Inventories – behavioral character traits
- Projective Tests – seemingly ambiguous - participant projects feelings into answers

Attitude Scale

- Pages 156 – 158 in text

A Question of Validity

Most important characteristic

Concerned with the appropriateness of interpretation

Tests are highly valid, moderately valid, or generally invalid for a particular interpretation and group

Three Types of Validity

- Content
- Criterion-related
- Construct

Content Validity

To what degree does a test measure the content? Requires both item and sampling validity

Item validity—to what degree does individual items measure content?

Sample validity—to what degree does test sample represent total content?

Criterion-Related

- Concurrent validity—relationship of test scores from tests administered at same time
- Predictive validity—relationship of test scores on test administered at different times: one now and one in the future
- Single group must take both tests

Construct Validity

- Most important form of validity
- Content and criterion-related tests for validity can be used
- Construct validity seeks to understand if the underlying variable is actually being measured
- Confirmatory and disconfirmatory evidence is reported

Diminishing Factors of Validity

- Unclear directions
- Inappropriate vocabulary
- Subjective scoring
- Administrative procedures

Reliability

- Consistency of measure
- Expressed numerically (coefficient 0.0 to 1.0)
- Correlation is the method to determine reliability

Approaches to Reliability

- Stability
- Equivalence
- Equivalence and stability
- Internal consistency
- Scorer/rater

Stability

- Expressed by correlating scores of one test of one group over time

Equivalence

- Two forms of one test given to single group

Equivalence and Stability

- Correlation of scores of one group taking similar tests taken at different times

Internal Consistency

- Measures the consistency among the items of a particular test

Scorer/Rater Reliability

- Measures how different raters or scorers score (grade) the same test

Forms of Reliability

- Split-half
- Kuder-Richardson
- Cronbach's alpha
- Spearman-Brown

Standard Error of Measure

- Estimates how often test score errors can occur
- High standard of error—high reliability
- Estimates difference between obtained score and true score
- Big difference in above—indicates low reliability

Choosing a Test

Consider

- Validity
- Reliability
- Ease of test
- Self-developed
  - Pretest
  - Revise
  - Pretest
  - Determine validity, reliability, and feasibility