Today’s Topics

- Data collection
- Measuring instruments
  - Terminology
  - Interpreting data
  - Types of instruments
- Technical issues
  - Validity
  - Reliability
- Selection of a test

Data Collection

- It’s all based on data
  - Scientific and disciplined inquiry requires the collection, analysis, and interpretation of data
  - Data – the pieces of information that are collected to examine the research topic
Data Collection: Terminology

- Data are often measurements of a **construct**
  - **Constructs** – abstractions that cannot be observed directly but are helpful when trying to explain behavior
    - Intelligence
    - Teacher effectiveness
    - Self-esteem

- **Operational definition** – specifies the specific tests/measures used to measure the construct of interest
  - Intelligence = standard scores on the Wechsler IQ test
  - Teaching Effectiveness = scores on the Virgilio Teacher Effectiveness Inventory
  - Self-esteem = scores on the Tennessee Self-Concept Scale

- **Variable** – a construct that has been operationalized

Data Collection: Variables

- Variables can be categorized as:
  1. Categorical or Continuous
  2. Independent or Dependent
Data Collection: Variables

Categorical or Continuous
- Defined from the type of data which represent them
- Categorical variables
  - reflect nominal scales
  - Gender (Male vs. Female)
  - SES (Low, Middle, & High)
  - Grade (1st graders, 2nd graders, etc.)
- Continuous variables
  - reflect ordinal, interval or ratio scale data
  - Academic achievement (score on the WJAT-II Test of Achievement)
  - Intelligence (IQ score on the WISC-IV)
  - Depression (Score on the Children's Depression Inventory)

Data Collection: Variables

Independent or dependent
- based on research question & design
- Independent variables (IV)
  - Variables thought to be the cause of a phenomenon under study
  - Often have several levels
    - Ex) VOEFS vs TUVEZ
      - Often have several levels
      - Ex) Reading Instruction with three levels (small group vs. large group vs. individual)
- Dependent variables (DV) are those that are affected by an independent variable(s)
  - Often measured by a test
    - Reading Test scores, Intelligence Test Scores, etc.
- Ex) Hypothesis: Vaccination causes autism
  - IV = Vaccination (two levels: vaccinated & not vaccinated)
  - DV = Number autism-like behaviors (Gilliam Autism Rating Scale score)

Data Collection: Example

- We want to study the effects of small and large group reading instruction on the reading achievement of second graders
  - Operational Definitions
    - Small Group Reading Instruction = 45 minutes of Instruction delivered in groups of 3 students
    - Large Group Reading Instruction = 45 minutes of Instruction delivered in groups of 10 students
    - Reading Achievement = Scores on the Woodcock Reading Mastery Test (WRMT)
Group Exercise

RESEARCH QUESTIONS:
▶ Are there differences between rural and urban children’s attitudes regarding diversity?
▶ Is there a relationship between post-secondary schooling and social competence?
▶ How will learners enrolled in an intensive summer math program achieve in math compared to those who are not enrolled in the program?
▶ What are the academic variables that account for a successful college experience?
▶ Is there a relationship between teachers’ training and job satisfaction?
▶ What characteristics of a school contribute to children’s attitude toward school?

EACH GROUP WILL:
▶ Select a Question
▶ Identify the dependent and independent variables.
▶ Develop the research question by operationally defining constructs within the dependent and independent variables.

Measurement Instruments

▶ Important terms (continued)
  ▪ Cognitive tests – examining subject’s thoughts and thought processes
  ▪ Affective tests – examining subject’s feelings, interests, attitudes, beliefs, etc.
  ▪ Achievement tests – examining subject’s reading, writing, or math skills
  ▪ Standardized tests – tests that are administered, scored, and interpreted in a consistent manner

Measurement Instruments:

▶ You can collect 4 types of data from measurement instruments
  ▪ Nominal – categories
    ▪ Gender, ethnicity, etc.
  ▪ Ordinal – ordered categories
    ▪ Rank in class, order of finish, etc.
    ▪ Don’t know the distance between positions. How much time passed between the race winner and runner up?
  ▪ Interval – equal intervals
    ▪ IQ scores, attitude scores, etc.
    ▪ The difference between IQ scores of 70 & 80 is the same as between IQ scores of 100 & 110.
    ▪ No absolute zero (A person with an IQ of 120 is not twice as smart as a person with an IQ of 60)
  ▪ Ratio – absolute zero
    ▪ Time, height, weight, etc.
    ▪ Allows direct comparisons between individuals on trait (a 4 ft. stick is twice as tall as a 2 ft. stick)
Measurement Instruments

- Interpreting data from measurement instruments
  - Raw scores – the actual score made on a test
  - Standard scores – statistical transformations of raw scores
    - Standard Scores
    - Z-scores
    - T-scores
    - Percentiles

Characteristics of a Normal Distribution

- Interpreting data (continued)
  - Norm-referenced – scores are interpreted relative to the scores of others taking the test
  - Criterion-referenced – scores are interpreted relative to a predetermined level of performance
  - Self-referenced – scores are interpreted relative to changes over time
Measurement Instruments

- Potential problems with measurement instruments
  - Bias – distortions of a respondent's performance or responses based on ethnicity, race, gender, language, etc.
  - Responses to affective test items
    - Socially acceptable responses
    - Accuracy of responses
  - Problems inherent in the use of self-report measures and the use of projective tests

Evaluating Tests

- What makes for a good test?
  - Reliability
    - The test is a good measurement tool . . . of whatever it's measuring
    - The specific construct of interest is not relevant
  - Validity
    - The test accurately measures the specific construct of interest

Test Reliability

- Reliability: A test’s consistency in measuring a specific trait or ability
  - Four types of reliability
    - Test-retest reliability (Stability): An index of a test's stability over time
    - Alternate form reliability: An index of consistency between different versions of a test
    - Internal consistency (split-half reliability): The extent to which all questions within in test measure the same thing
    - Inter-rater reliability: The extent to which different examiners produce similar results with a test
  - Listed in test manuals and expressed as a reliability coefficient (r)
    - r values range from 0.00 to 1.00
    - Higher r values indicate higher reliability
    - r values should be around .80
Test Validity

- **Validity**: The extent to which a test measures what it claims to measure
  - Revolves around two broad questions:
    - What does a test measure?
    - How well does it measure it?
  - Is directly related to the purpose of a test

Test Validity: Content Validity

- **Content Validity**: the extent to which the items on a test are representative of the constructs it claims to measure
  - e.g., How thoroughly are you measuring the desired construct or trait?
  - Does the test measure the domain of interest?
  - Are the test questions appropriate?
  - Does the test contain sufficient information to appropriately cover what it is supposed to measure?
  - What is the level of mastery at which the content is being assessed?
Test Validity: Construct Validity

- **Construct Validity:** the extent to which a test measures a psychological construct or trait (e.g., Does your test actually measure the desired construct?)

Test Validity: Criterion-Related Validity

- **Criterion-Related Validity**
  - The relationship between test scored and some type of outcome
  - Other outcomes can include ratings, classifications, or other test scores
    - **Concurrent Validity:**
      - The extent to which a test is related to other assessments of the same construct
      - Will a child who earns good grades in math also score highly on a test measuring math skills?
    - **Predictive Validity:**
      - The extent to which a test predicts future outcomes on a related criteria
      - Does a reading test given at the start of the school year predict reading performance at the end of the year?

Test Validity: Predictive Utility

- **Predictive Utility**
  - The extent to which a test agrees with a criterion measure in classifying individuals a to their membership in a category
  - **Example:**
    - How often does a behavior rating scale correctly identify kids diagnosed with ADHD
Factors Affecting Validity

- Overly difficult and complex sentence structure
- Inconsistent and subjective scoring
- Untaught items (achievement tests)
- Failure to follow standardized administration procedures
- Cheating by the participants or someone teaching to the test items

Selecting a Test: Issues to Consider

- Psychometric properties
  - Validity
  - Reliability
  - Length of test
  - Scoring and score interpretation
- Non-psychometric issues
  - Cost
  - Administrative time
  - Objections to content by parents or others
  - Duplication of testing
Designing Tests: Issues to Consider

- Get help from others with experience developing tests
- Item writing guidelines
  - Avoid ambiguous and confusing wording and sentence structure
  - Use appropriate vocabulary
  - Write items that have only one correct answer
  - Give information about the nature of the desired answer
  - Do not provide clues to the correct answer

Resources about Tests

- Sources of test information
  - Mental Measurement Yearbooks (MMY)
    - Provides factual information on all known tests
    - Provides objective test reviews
  - Comprehensive bibliography for specific tests
  - Indices: titles, acronyms, subject, publishers, developers
  - Buros Institute

Resources about Tests (continued)

- Tests in Print
  - Bibliography of all known commercially produced tests currently available
  - Very useful to determine availability
  - Tests in Print
Resources about Tests

▶ Sources (continued)
  • ETS Test Collection
    ▶ Published and unpublished tests
    ▶ Includes test title, author, publication date, target population, publisher, and description of purpose
    ▶ Annotated bibliographies on achievement, aptitude, attitude and interests, personality, sensory motor, special populations, vocational/occupational, and miscellaneous

Resources about Tests

▶ Sources (continued)
  • ERIC/AE Test Locator
    ▶ Search for citations about a particular instrument
    ▶ Search for names and addresses of test publishers

Resources about Tests

▶ Sources (continued)
  • Professional journals
  • Test publishers and distributors