Lecture 7   Psychological Testing and Measurement
Sunthud Pornprasertmanit

Test Utility

Classroom Assignment

Case 12.2

- Examine: Item difficulty, Item discrimination and Distracters Statistic

Test Utility

Lecture 7   Psychological Testing and Measurement
Sunthud Pornprasertmanit

Test Utility

- How the test can be used in the real world situation?
- In some situations, the well-validated test may not be utilized.
- However, in some situations, the adequate-validated test may be very useful.

Test Utility

- Test users frequently use the test to be decide for solving some questions?
  - Selection
  - Diagnosis
  - Screening
  - License certifying

Test Utility

- How to decide
  - Pass/Fail
  - Grading
- Which rule
  - Norm-referenced
  - Criterion-referenced

Test Utility

- Norm-referenced decision based on individual differences.
- People who selected must be one of the best in the reference group (applicants).
- However, the best people may not be qualified.

Test Utility

- Method in grading or selecting based on
  - Best People
  - Percentile
  - *z*-score

Test Utility

- Criterion-referenced decision based on standard.
- How can we set the standard?
- Pass/Fail selection based on minimal acceptance criterion.

Test Utility

Minimal Acceptance Criterion is the qualification of Minimal Acceptance Person (MCP)

- Judgmental Method
- Judgmental/Empirical Method
- Empirical/Judgmental Method
- Empirical Method

Test Utility

- Judgmental Method
  - Angoff: rated 0-100 % in each item that represent MCP corrected and then average
  - Nedelsky: eliminate distracters that MCP did not answer and the average percent guessing
  - Ebel: divided in difficulty and relevance and then rated percent MCP corrected in each cell

Test Utility

- Judgmental/Empirical Method
- Delphi Technique
- Use techniques in judgment method
- Then, pooled by moderator and adapted by experts as a iterative process
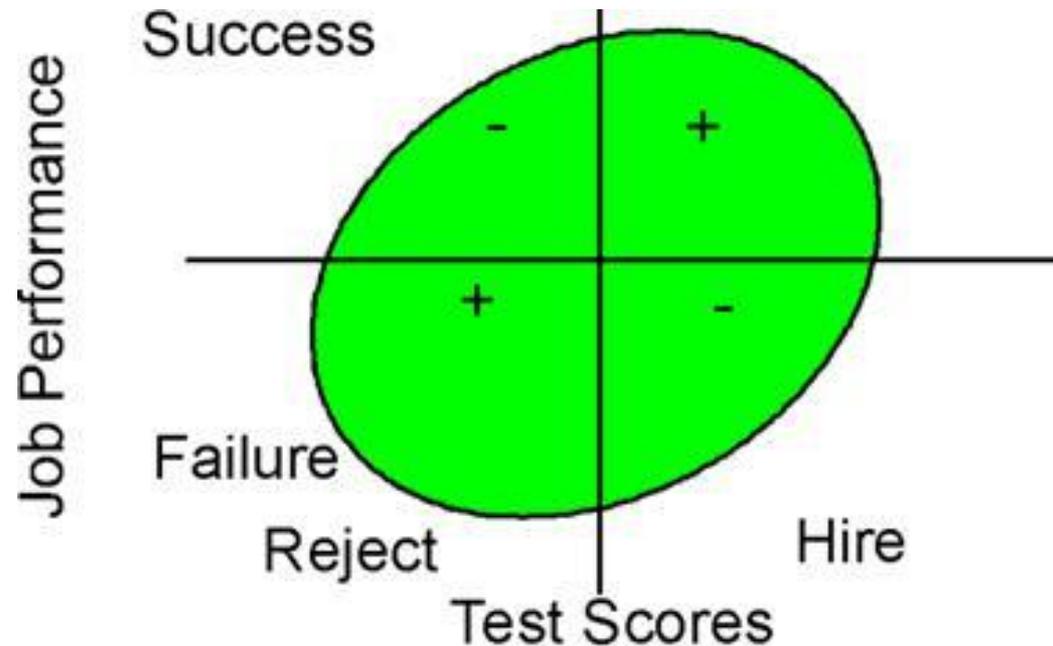
Test Utility

- Empirical/Judgmental Method
- Known group
- Cutoff score is the score that maximal discriminate between achiever and non-achiever.

Test Utility

- Empirical Method
- Like using Norm-referenced decision
- Determine the cutoff score by the reference group.

Test Utility

- Result of decision
  - Correct acceptance
  - Correct Rejection
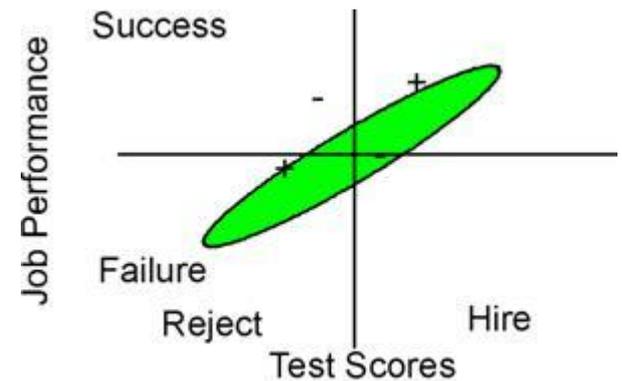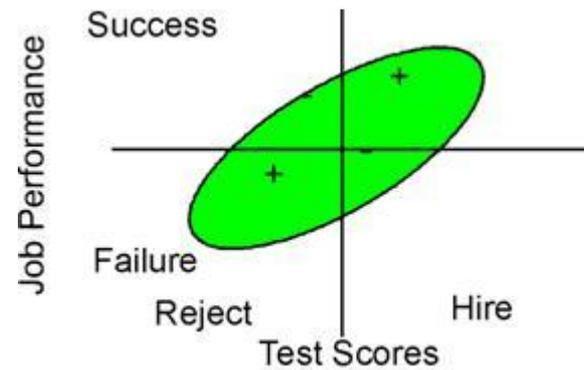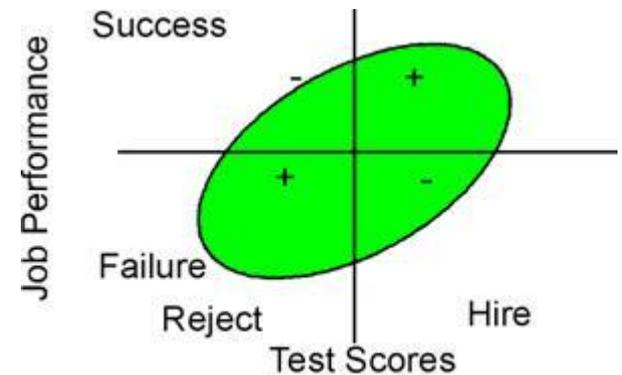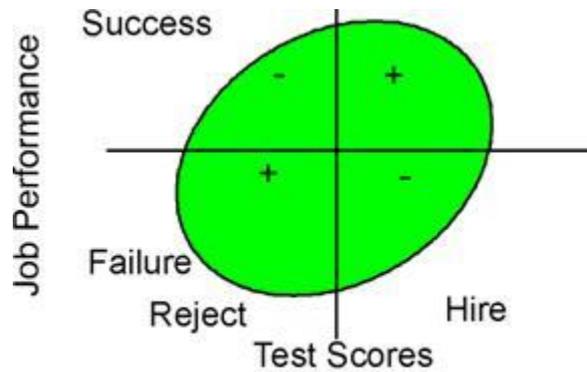 - False Positive
  - False Negative

Test Utility

- How validity link with result of decision
- Taylor and Russell (1939) table
- Three factors that affect the decision
  - Validity (Correlation)
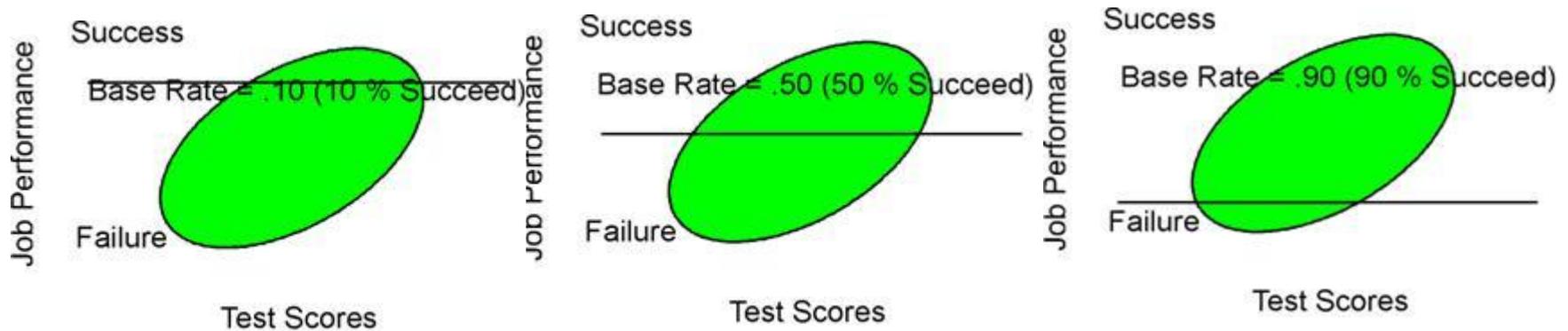 - Base rate (Achievers within applicants)
  - Selection Ratio (Cutoff)

Test Utility

- Validity

Test Utility

- Base Rate

Test Utility

- Selection Ratio
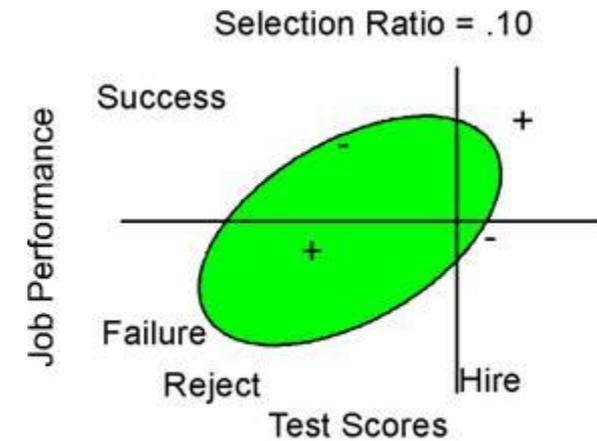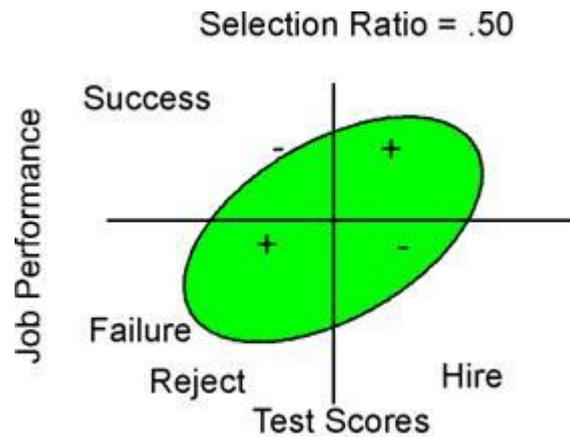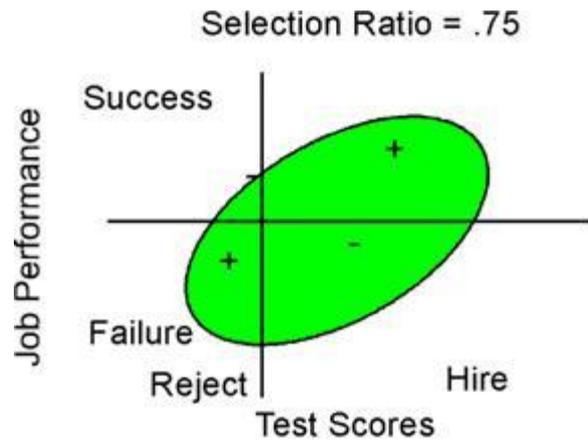
Test Utility

- When did not use the test: (A+B)/(A+B+C+D)
- When use test: A/(A+D)
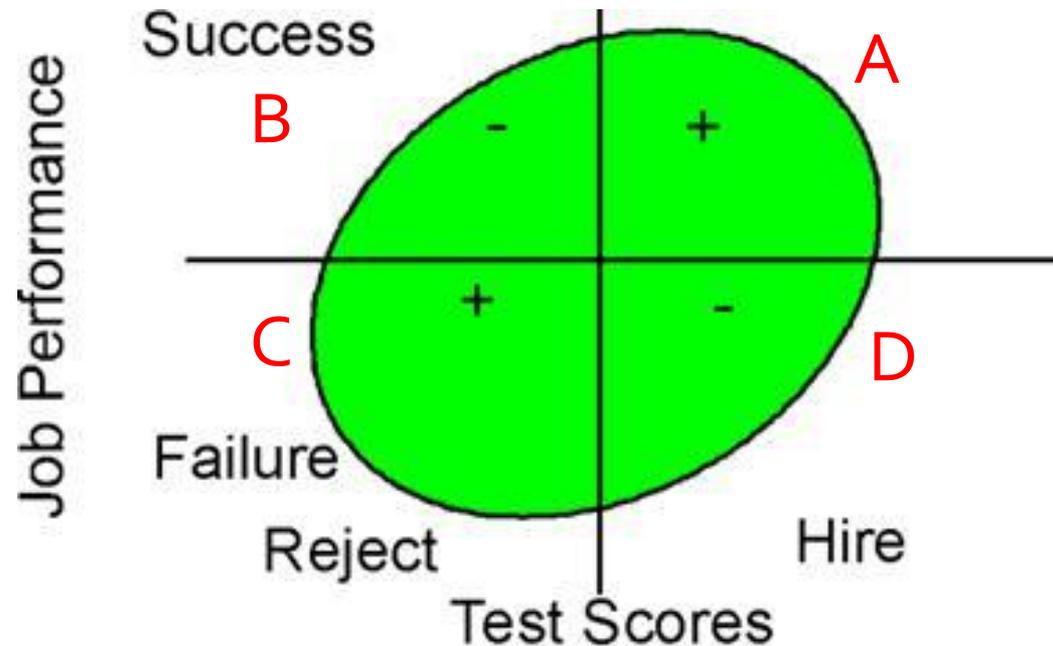
Test Utility

- Find A/(A+D) by table in Taylor and Russell (1939) or **Employment Statistics Homepage**
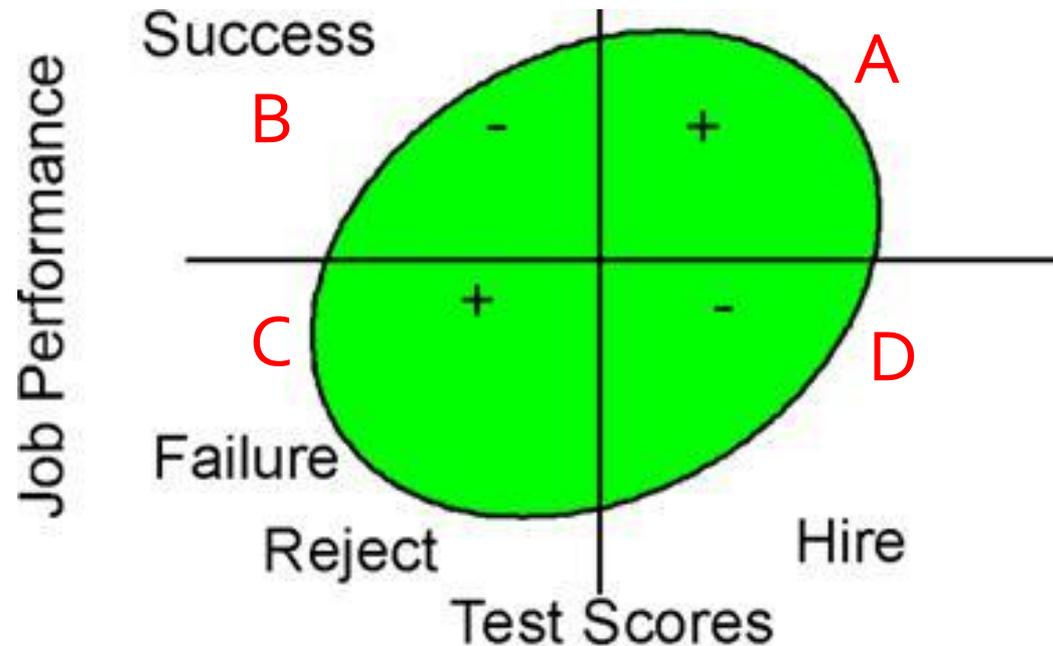
Exercise

Computer Exercise 5.1, 7.3, 8.4, 12.2

Sales.sav

- Reliability (Alpha)
- Criterion-related validity (Correlation b/w cognitive test and sales performance)
- Item Difficulty (Percent Correction)
- Corrected Item-total Correlation
- Item Discrimination with external criterion (Regress sales performance on items)

Class Assignment

Homework

- Finding the problem of EQ definition
- Find the ways to validate the EQ measures as much as possible from the article
- Which concepts are new for you in this article

Class Assignment

Enhance the construct validity evidences
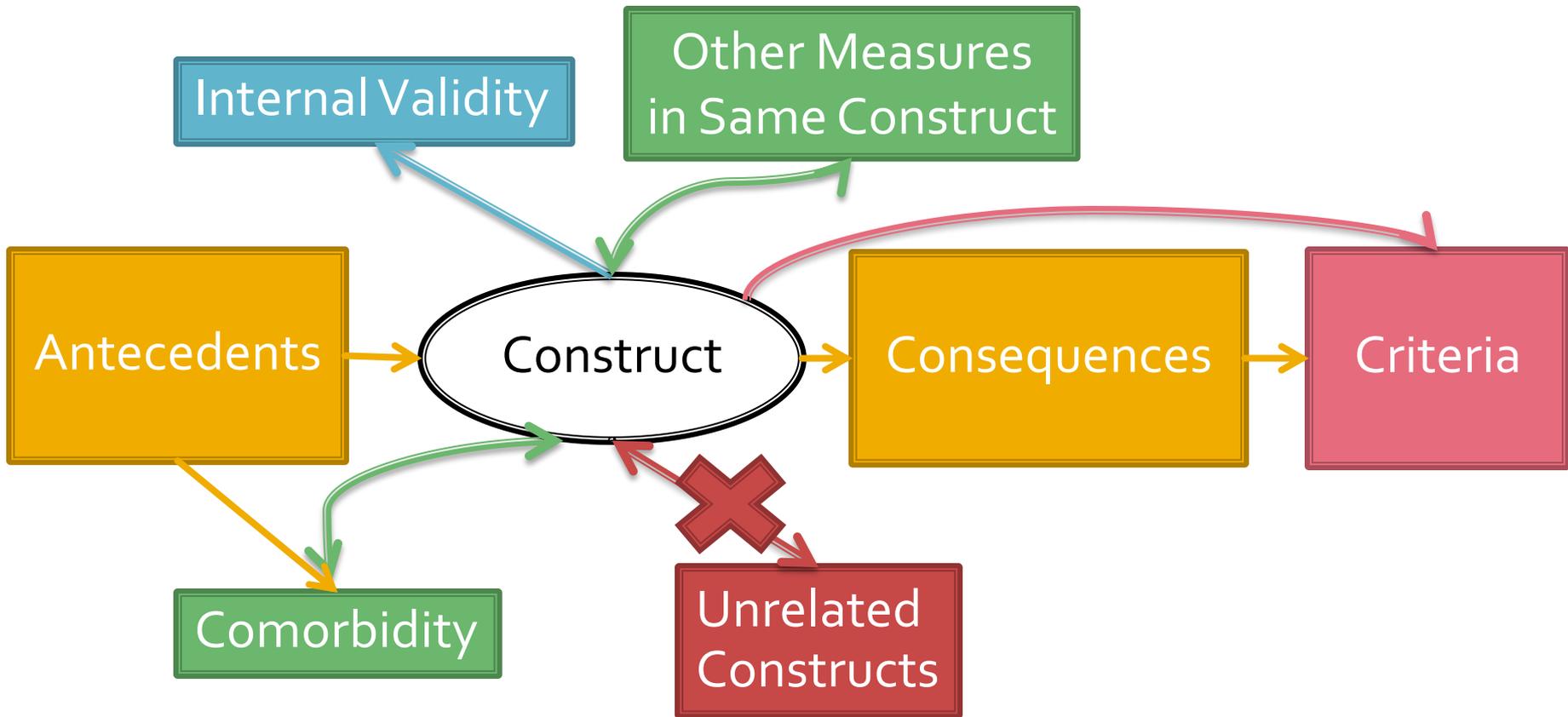
Classroom Assignment

- What does emotional intelligence mean?
  - Abilities or personalities
- How do you know?
  - 4 factors?
  - Is it exhaustive? Are factors representative of EI?
- Is it true?
  - Emotional intelligence as a construct?
- Can it be explained otherwise?
  - Consideration, consider other mind,

Classroom Assignment

- Validity defines by four previous questions.

Classroom Assignment

- Developing 20 multiple choice items for screening psychological test users
- Thinking how to build the test to achieve desired reliability and validity
- Then use Judgment (Angoff) and Judgment/Empirical Method (Delphi)

Classroom Assignment

Interview 5 people (different faculties):
1) how to prepare for testing
2) Find the strategies of test takers for coaching and gaining the test sophistication
3) physical and psychological aspects before testing
4) thinking of cheating, how to cheating and why cheating
5) Test strategies (e.g. method for guessing)

Test Use

Lecture 7 Psychological Testing and Measurement
Sunthud Pornprasertmanit

The Role of Test User

- Test Selection
- Test Administration
- Test Scoring
- Test Interpretation
- Test Confidentiality

Qualification for Users of Psychological Tests

- Their knowledge and skills in
    - Psychometric principle and statistics
    - Selection of tests in light of their technical qualities, the purpose for which they will be used, and the characteristics of examinees
    - Procedures for administering and scoring tests, as well as for interpreting, reporting, and safeguarding their results
    - All matters relevant to the context and purpose of test use

Qualification for Users of Psychological Tests

- Test users must be trained and supervised in the overall skills and specific skills in some tests.

Test Selection

- Before choosing tests, consider following questions

    - What kind of information do I seek to gain from testing?

    - How will this information be used?

    - How much of the information I seek is already available from other sources?

Test Selection

- Before choosing tests, consider following questions
  - What other tools might be used to gather the information I seek?
  - What are the advantages of using tests instead of, or in addition to, other sources of information
  - What are the disadvantages or the costs in time, effort, and money of using tests?

Test Selection

- The most significant advantages that psychological tests offer pertain to their characteristic efficiency and objectivity.

  - Efficiency: timely and cost-effective manner.

  - Objectivity: The tests depend on normative or criterion-based frame of reference.

Other Sources of Information

- Biographical Data
- Interview
- Assessment Center
- Observation
- Letter of Recommendations
- Polygraph

Test Administration

- Environment
- The test takers aspect
- The examiner aspect
- Interaction between the Examiner and the Subject
- The technology aspect

The Test Environment

- The most important principle to follow in preparing the environment in which testing will occur is to anticipate and remove any potential sources of distraction.

  - Light, Ventilation, Seating, Space, Noise, Other Stimuli, Sign for Direction
  - Trying to prevent cheating behavior
  - For individual test, presence of others in testing room

The Test Takers

- Test Takers' Variables
  - Motivation and anxiety can greatly affect test scores.
  - Test anxiety: students who suffer from it have difficulty focusing attention on the test items and are distracted by thoughts such as "I am not doing well" or "I am running out of time"
  - Test anxiety appears to have three components: worry, emotionality, and lack of self-confidence

The Test Takers

- Test Takers' Variables
  - Many variations in health status affect performance in behavior and in thinking, such as illness, medical drugs.
  - Some researchers have debated whether normal hormonal variations affect test performance.

The Test Takers

- Test Takers' Variables
  - Social desirability: trying to fake good or bad
  - May or may not be conscious
  - Validity scales designed to detect various types of attempts at impression management or response sets such as defensiveness are embedded in a number of personality inventories

Response Biases

- Guessing
  - Correction of guessing
  - Personality that affect the response style in correction of guessing
  - Computer adaptive test

Response Biases

- Rating Error
  - Central tendency effect
  - Leniency effect
  - Severity effect
- Attitude Rating Error
  - Acquiescence bias
  - Nonacquiescence bias

Response Biases

- Response Biases
  - Faking good situation: Impression Management
  - Faking bad situation: Criminal Justice
- Response Styles

Response Biases

- How to prevent
  - Clear instruction
  - Anonymity (if possible)
  - Forced-choice???
  - Changing methods (e.g. unobtrusive measure)
  - Lie scale or social desirable scale???

Effects of Training on Test Performance

- Is improvement is limited to specific items included in the test or it extends to the broader behavior domain that the test is designed to assess?
- A test score is invalidated only when raises the score without affecting the behavior domain that the test is designed to measure.
- Says "Test Coaching"

Effects of Training on Test Performance

- Coaching is designed to develop highly restricted skills that may be of little use in life activities.
- Coaching may be available to some test takers and not to others, in tends to introduce individual differences in narrowly defined test-taking skills.

Effects of Training on Test Performance

- Individual with deficient educational backgrounds are more likely to benefit from special coaching.
- The closer the resemblance between test content and coaching material, the greater will be the improvement in test scores.

Effects of Training on Test Performance

- Test Sophistication:
- Significant mean gains in alternate form after the original form
- Ranging from one day to three years, both in same form and alternate form.

Effects of Training on Test Performance

- Test Sophistication:
- Experience test takers are higher in performance than one who is taking her or his first test.
- Such as initial feeling of strangeness, more self-confidence, better test-taking attitudes

Effects of Training on Test Performance

- For computer administered testing, attention should be given to the test taker's familiarity with this form of test administration.
- Short orientation and practice sessions can be quite effective in equalizing test sophistication.

Cheating

- Test cheating affect the meaning of the score
- 60 % of the CU bachelor students has cheated (both give and take)
- Factors affect cheating
    - Demographic Variable (e.g. sex, education level)
    - Administration
    - Test format
    - Positive and Negative Expectancies
    - Personality (Sensation seeking)

Cheating

- Even typical performance test may be affected by copying other people responses.

The Examiner

- Training of Test Administrators
    - Different assessment procedures require different levels of training.
    - Especially, in individual test
    - In a study of 22 graduate students, there were numerous errors in scoring the test (WISC), with no improvement over five practice administrations.

The Examiner

- Informed Consent: either orally or in writing.
- contains a suitable explanation of nature and purpose of the evaluation,
- as well as information concerning confidentiality limits and how the security of test results will be maintained.

The Examiner and the Test Taker

- Bias provided by examiners
  - Relationship b/w examiner and test taker
  - Interviewer effect
  - Race
  - Language
  - Expectancy Effect
  - Reinforcement

The Examiner and the Test Taker

- The relationship between examiner and test taker
  - Feldman & Sullivan (1960): first- through seventh-grade children were given the WISC under one of two conditions (rapport, neutral)
  - Fuchs & Fuchs (1986) take meta-analysis on 21 studies found that test performance was approximately .28 SD higher when the examiner was familiar with the test taker than when not.

The Examiner and the Test Taker

- Interviewer Effect
  - Respondent may give the response that they perceive to be expected by the interviewer.
  - Moun (1998): assess symptoms of anxiety and depression in one of three conditions: face-to-face, telephone and self-administered quesionnaire
  - Self-administered questionnaires = 2/3 personally interviewed cases.

The Examiner and the Test Taker

- The Race of the Tester
  - Some groups feel that their children should not be tested by anyone except a member of their own race.
  - Sattler (2002, 2004) little evidence that the race of the examiner in intelligence test scores.
  - The race of the examiner affects test scores in some situations, such as discretion about the use of the tests, paraprofessional.
  - Stereotype Threat (Steele, 1997)

The Examiner and the Test Taker

- Language of Test Taker
  - Some test are inappropriate for people whose knowledge of the language is questionable.
  - Translating tests is difficult, and it cannot be assumed that the validity and reliability of the translation are comparable to the English version.
  - Evidence for test comparability across language should be available.

The Examiner and the Test Taker

- Expectancy Effect
  - Expectancy effect exists in some but not all situations.
  - Two aspects of expectancy effect relate to the use of standardized tests.
    - Nonverbal communication to test taker
    - Scoring of examiner
  - Inconclusive results of expectancy effect b/c very small sample size (Sattler, 1988)

The Examiner and the Test Taker

- Effect of Reinforcing Responses
  - An inconsistent use of feedback can damage the reliability and validity of test scores.
  - Reward (such as tokens, praise) can significantly affect test performance, especially cultural relevant reward.
  - However, it is not clear that reward improve or harm scores.

The Examiner and the Test Taker

- Effect of Reinforcing Responses
  - The way an interviewer responds affects the content of responses in interview studies.
  - Because different examiners make different responses, one cannot ensure that the advantages resulting from reinforcement will be the same for all people.
  - As a result, most test manuals and interviewer guides insist that no feedback be given.

The Examiner and the Test Taker

- Effect of Reinforcing Responses
  - A good test manual gives the test examiner instructions that include the exact words to be read to the test takers.
  - A few occasions do require deviation from standardized testing procedures, such as for disabilities, elderly.

The Technology

- Computer-Assisted Test Administration
  - Advantages that computer offer
    - Excellence of standardization
    - Individually tailored sequential administration
    - Precision of timing responses
    - Release of human testers for other duties
    - Patience
    - Control of bias

The Technology

- Computer-Assisted Test Administration
  - Advantages that computer offer
    - Automatic recording of test responses
    - Easy of application of complicated psychometric issues and the integration of testing and cognitive psychology
    - Items can be given in any order
    - Prevents test takers form looking ahead at other sections of the test or going back to sections already completed
    - Computer test reported

The Technology

- Computer-Assisted Test Administration
  - Some people, though, uneasy interacting with computers, or suffer from "keyboard phobia."
  - Comparisons of test scores have not tended to show large differences between computer-assisted and paper-and-pencil tests.
  - Test takers find interactions with computers more enjoyable than paper-and-pencil tests.

The Technology

- Computer-Assisted Test Administration

  - Type of tests are prepared for computer administration, independent reliability and validity studies will be needed.

  - In clinical judgment, like an inexperienced psychologist, computerized reports may cause harm.

Test Scoring

- In scoring both by hand and by scanner, there is possibility of errors.
- Good testing practice requires double-checking the accuracy of all computations and equivalent scores obtained from tables.

Test Scoring

- When required rating or subjective judgment, the test examiner must be trained to rate in valid and reliable manner.
- Scorer reliability coefficient must not be below .90.

Test Score Interpretation and Reporting

- Test score interpretation: Test takers can either derive the greatest benefit, or incur the greatest harm.
- Data from multiple sources are often necessary

Test Score Interpretation and Reporting

- In interpretation, it requires knowledge about the context and specific areas of human behavior relevant to issue in question.
- Consider the consequences of the decisions that are made.

Test Score Interpretation and Reporting

- In actual practice, contextual issues are frequently ignored,
- test scores often become the major or even the sole determining factor in decision-making.
- One way to counteract this problem is to understand and appreciate the implications of the difference between psychological testing and assessment.

Test Score Interpretation and Reporting

- The most basic guideline to follow in communicating test results is to provide
  - the information derived from test scores,
  - including its limitations, in language that the recipient can understand.

Test Score Interpretation and Reporting

- Don't
  - Report numerical scores
  - Assigning labels

Test Score Interpretation and Reporting

- Do
  - At a minimum: what the test covers, the meaning of scores,
  - the limitations on the precision of scores that derive from measurement error,
  - common misinterpretation to which particular scores,
  - the way in which test results may or will be used
  - At its best: use multiple sources

Test Confidentiality

- The security of test data,
- whether they consist of individually identifiable records, scores, and reports or of test materials themselves,
- is a primary responsibility of the test users and institutions who control access to such data.

Next Lecture Developing PT

Lecture 5 Psychological Testing and Measurement
Sunthud Pornprasertmanit

Developing PT

Example
- Drinking and Driving Positive Expectancies
- Mindfulness Scale
- Team Role Test