NORMS AND RELIABILITY

I. NORMS (Comparing Frame of Reference)

    A.   Norm Group

        1. Representative of the population

        2. Test Standardization - allows for

           publishing derived scores

        3. Raw Scores - meaningless in isolation

        4. Non -referenced tests

    B.   Essential Statistical Concepts

        1. Frequency Distribution

           - equal sized class intervals

           - Histogram (graphic representation)

           - Frequency Polygon (single point)

        2. Central Tendency

           - mean

           - median

           - mode

3. Measures of Variability

- Standard Deviation P. 61

    i)   **S, SD - square root of variance**
    ii) **degree of dispersion in a group of scores**

- Normal Distribution

    i) **useful mathematical features for statistical investigation.**
    ii) **mathematical precision**
    iii) **arises spontaneously in nature.**

- Skewness

    positive - low end
    negative - high end

C. Raw Score Transformation

1. Percentiles and percentile rank

2. Standard Scores - **distance from the mean (p. 65) in standard deviation units (a "z" score)**

3. T-scores and other standardized scores
    T-score - mean 50   SD 10
    Standardized scores eliminate - negative; fractions; 0's.

4. Normalizing Standard Scores

    - percentiles to standard scores

5. Stanine Stem - C-scale

6. Figure 3.8 (p.69)

D. Selecting a Norm Group

1. Random Sampling
2. Stratified Random Sampling
3. Age and Grade Norms
4. Local and Subgroup Norms
   (used continuously)
5. Expectancy Tables
   - test scores and expected outcomes
   (p.72 - Table 3.6)

E. Criterion - Referenced Tests

1. Compare examinees accomplishments to a predefined performance standard. ie. what they can do (mastery)

2. Table 3.7 (p. 74) Comparison of Norm vs. Criterion Referenced Tests

RELIABILITY

I.   Sources of Measurement Error

   A. $X = T = e$ $e = X - T$

      (true score is never known)

   B.   Item Selection   Unsystematic

   C.   Test Administration   Measurement

   D.   Test Scoring   Error

   E.   Systematic Measurement Error

      $X = T + e_s + e_u$

II.   Measurement Error and Reliability (p. 81)

   A.   Randomness

   B.   Mean error of measurement = 0

   C.   True scores and errors are uncorrelated

   D.   Errors on different tests are uncorrelated

III.   The Reliability Coefficient

   A.   "The ratio of true score variance too the total

      of test scores" $r_{xx} = \dfrac{6^2T}{6^2T + 6^2e}$

   B.   1.0 to 0.0

   C.   The Correlation Coefficient (-1.00 to 1.00) P. 82

IV. Reliability as Temporal Stability
- A. Test- Retest Reliability
- B. Alternate - Forms Reliability
  - (Immediate / Delayed)

V. Reliability as Internal Consistency
- A. Split - Half Reliability
- B. The Spearman - Brown Formula
- C. Coefficient Alpha
- D. The Kuder-Richardson Estimate of Reliability
  - (KR-20)
- E. Interscorer Reliability

VI. Special Circumstances
- A. Unstable Characteristics
- B. Speed and Power Tests
  speed: need test - retest reliability
- C. Restriction of Range
- D. Criterion - referenced Tests
- E. Interpretation of Reliability Coefficients (.70 to .95)

VII. Standard Error of Measurement
- A. Std Deviation of the distribution of obtained scores.
- B. $SEM = SD\sqrt{1-r}$ (p. 92)
- C. Confidence Intervals
- D. Standard Error of the Difference

<u>VALIDITY</u>   <u>(Accuracy)</u>

I.   Introduction

- Reliability (expressed as a correlation coefficient) is a necessary but not a sufficient precursor to validity.

- Validity is a developmental process that begins with test construction and continues indefinitely.

- Page 97 Case Exhibit 4.1

- A test is valid to the extent that inferences made from it are appropriate, useful and meaningful.

- Not defined by a single statistic but on an evolving continuum (weak to acceptable to strong).

- Three categories of validity: content, criterion-related, construct.

II.   Content Validity

    A.   The degree to which the test is representative of the universe of behavior the test was designed to sample.

    B.   Useful for areas where a great deal is known about the variable we wish to measure (especially achievement tests).

    C.   Difficult to ensure when the test measures an ill-defined trait.

    D.   Quantification of Content Validity (P. 99 - 100)

    E.   Face Validity

III. Criterion-Related Validity

    A.   Validity Coefficient

        1. Constrained by reliability of test and criteria (p. 101)

        2. A criterion measure must be appropriate to the test under investigation.

3. Must be free from contamination

    (criteria contamination)

B. Concurrent Validity (eg. diagnostic tests)

    - Correlated with other tests

    - with non-test behavior in the world

    - must measure the same construct as the criteria

C. Predictive Validity

    - employment tests

    - requires a regression equation

    - validity coefficient

    - Standard Error of Estimate (p. 103)

      (margin of prediction error caused by the

        imperfect validity of the test.)

      Similar to the SEM

B. Decision Theory

- P. 104 (Figure 4.4)

   Measurement in the Service of Decision Making

   Selection Ration (0.0 - 1.0)

   False positives, False negatives (misses/hits)

   Taylor-Russell Table P. 107

- Base Rate

- Even law validity test are beneficial

IV. Construct Validity

A theoretical, intangible quality or trait in which individuals differ. (eg. leadership ability)

A. Test homogeneity

B. Developmental Changes

C. Theory consistent with group differences

D. Convergent Validity

E. Discriminant Validity

F. Multitrait / multimethod matrix (Table 4.4 p. 112)

G. Factor Analysis (Table 4.5 p. 113)

V. Extravalidity Concern

- Unintended side effects

- technical vs. functionalist view

- constructive consequence / guard against adverse

outcomes

VI The Test Construction Process

Figure 4.5 Page 119

Nominal, Ordinal, Interval, Ratio Scales (Fig. 4.6)

Chapter 5

Intelligence

I. **Definition (p. 165)**

- **Wechsler (1939)**

- **The capacity to learn from experience**

- **The capacity to adapt to one's enviroment**

- **Experts: 1. Verbal Intelligence**

        2. Problem-solving

        3, Practical Intelligence

- **Lay person: 1. Practical problem solving**

        2. Verbal ability

        3. Social capability

- **Table 5.1 (167)**

- **Factor Analysis (p. 170-173)**

II. **Theories of Intelligence**

A. **Galton – Keen Sensory Abilities**

B. **Spearman – g / s**

C. **Thurstone – primary mental abilities (p.178)**

Theories of Intelligence cont.

D. Cattell-Horn-Carroll (CHC) p. 179, fig. 5.6

   1 (g) – 8 (broad factors) – 70 (narrow factors)

   R. Cattell – fluid/ crystallized

E. Piaget – Table 5.6, p. 182

F. Guilford: Structure-of-Intellect model (p. 181)

   (Operations / Content / Products)

G. Simultaneous / Successive Processing (Luria)

H. Information Processing Theories of Intelligence

   (architectural / executive)

I. Biological Construct (EEG)

J. Gardner: Theory of Multiple Intelligence

K. Sternberg and Triarchic Theory of Intelligence

   (Table 5.7, p. 180)

III. Individual Tests of Intelligence and Achievement.

1. WAIS-III   {p. 198

2. WISC-IV   {table 5.8

3. Stanford-Binet V edition p. 212, fig. 5.15

    a) high and low end items

    b) religious diversity

4. Detroit Tests of Learning Aptitude – IV

    - needs empirical support.   p. 214, Table 5.11

5. KABC-II - Luria - CHC

    p. 215, fig. 5.15

6. KAIT – 11 to 85 CH F/C

7. KBIT II

8. Individual Tests of Achievement – p.222, Table 5.14

    - usually a discrepancy between general ability and

       specific achievement.

9. p. 226, fig. 5.16

10. p. 228, Table 5.16

Chapter 6: Group Tests

I. **Introduction**

A. **Group Tests vs. Individual Tests**
1. Multiple Choice vs. open-ended format
2. objective machine scoring vs. examiner scoring
3. group vs. individual administration
4. application in screening vs. remedial planning
5. huge vs. large standardization sample

B. **Risks of Group Tests**
1. Low scores due to poor motivation or difficulty following directions.
2. Invalid scores are not recognized as such by examiner.

II. **Group Tests of Ability**

A. **MAB-II (approximates WAIS-R)**

B. **Shipley Institute of Living Scale**

C. **Cognitive Abilities Test**

D. **Culture Fair Intelligence Test**

E. **Raven's Progressive Matrices**

III. **Multiple Aptitude Batteries**

    A.   **Differential Aptitude Test**

    B.   **General Aptitude Test Battery**

    C.   **Armed Services Vocational Aptitude Battery (ASVAB)**

IV. **Predicting College Performance**

    A.   **Scholastic Assessment Tests (SAT)**

    B.   **American College Test (ACT)**

V. **Postgraduate Selection Tests**

    A.   **Graduate Record Exam (GRE)**

    B.   **Medical College Admission Test (MCAT)**

    C.   **Law School Admission Test (LSAT)**

VI. **Educational Achievement Tests (p. 258)**

    A.   **Iowa Test of Basic Skills (ITBS)**

    B.   **Metropolitan Achievement Test**

    C.   **General Educational Development (GED)**

VII.   **Effects of Coaching on Aptitude Tests**

Chapter 6B: Test Bias and Controversies

I.   **Test Bias – an objective, empirical question; not a matter of personal judgment. (p. 276)**

II. **Test Fairness – an ethical question**

III. **Genetic and Environmental Determinants of Intelligence - Nature vs. Nurture (Heritability Index)**

   A. **Teratogenic Effects (Fetal Alcohol Syndrome)**

   B. **Environmental Toxins**

   C. **Racial Differences**

   D. **Age Changes**

   E. **Generational Changes – Flynn Effect**

Chapter 7A: Infant and Preschool Assessment

I. **Infant Ability**

    A. **Gesell Development Schedule**

    B. **Neonatal Behavioral Assessment Scale**

    C. **Ordinal Scales of Psychological Development**

    D. **Bayley III**

    E. **Additional Measures   p. 301, Table 7.1**

II. **Preschool Intelligence**

    A. **McCarthy Scales of Children's Abilities**

    B. **Differential Ability Scales**

    C. **WPPSI – III**

III. **Practical Utility**

    A. **Screening for Developmental Disabilities**

    B. **Fagan Test of Infant Intelligence**

IV. **Screening for School Readiness**

    A. **Developmental Indicators for the Assessment of Learning III (DIAL III)**

    B. **Denver II**

    C. **HOME (p. 314)**

Chapter 7 Supplement

A. **<u>Definition of Mental Retardation</u>**

Over 200 etiological factors leading to impaired efficacy to learn.

IQ 2 SD below x

before age 18

observable deficits in ability to function

B. **<u>Levels of Mental Retardation</u>**

1. **Profound**

    WAIS III < 25

2. **Severe**

    WAIS III 25 – 40

3. **Moderate**

    WAIS III 40-55

4. **Mild**

    WAIS III 55-69

5. **Borderline Range**

    WAIS III 70-79

Chapter 7B: Testing Persons with Disabilities

I. **Non-Language Tests**

    A. **Leiter International Performance Scale - Revised**

    B. **Human Figure Drawing Tests**

    C. **Hiskey-Nebraska Test of Learning Aptitude**

    D. **Test of Non-Verbal Intelligence – 3**

II. **Non-Reading and Motor-Deduced Tests**

    A. **Peabody Picture Vocabulary Test - III**

III. **Visual Impairments**

    A. **Haptic Intelligence Scale for the Blind**

    B. **Intelligence Test for Visually Impaired Children**

IV. **Hearing Impairment**

    A. **WAIS-III in ASL**

V. Mental Retardation

    A. Definition – p. 332

    B. Levels   p. 333, Table 7.12

    C. Adaptive Skills and Behavior

        1. Scales of Independent Behavior – Revised (Table 7.1)

        2. Independent Living Behavior Checklist

        3. Inventory for Client and Agency Planning

        4. Vineland Adaptive Behavior Scales

Theories of Personality

I.   Differences between People / Consistency within Individuals

    A.   **Psychoanalytical Theory**

       1. Id – pleasurable principle

       2. Ego – reality principle / defense mechanism

         (table 8.1, p. 345)

       3. Superego – guilt principle

    B.   **Type Theories**

       Type A, coronary prone (p. 347)

    C.   **Phenomenological Theories**

       Carl Rogers - Q-technique (p. 349)

       Self-theory / Self Concept

    D.   **Behavioral / Social Learning Theories**

       Locus of Control (I/E) - Rotter

       - source of things that happen to them

       Bandura - Self-efficacy

    E.   **Trait Conception**

       Trait – relatively enduring way in which one individual

       differs from another.

       1.   Cattell - 16 Source Traits

       2.   Eysenck - Two-dimensional (Fig. 8.1, p. 353)

       3.   Goldberg – 5 Factor Model (p. 352)

Projective Tests

Vague stimuli reveal underlying needs, motives, and conflicts.

I. **Association Techniques**

    A. **Rorschach Inkblot Test**

    B. **Holtzman Inkblot Technique – Schizophrenia**

II. **Completion Techniques**

    A. **Sentence Completion Technique**

    B. **Rotter Incomplete Sentence Blank**

    C. **Rosenzweig Picture Frustration Story - Research**

III. **Construction Techniques**

    A. **TAT – Henry Murray (Need-Press)**

        T-TAT: African Americans

        TEMAS: Hispanics

    B. **Picture Projective Test - Photo**

    C. **Children's Apperception Test**

IV. **Expression Techniques**

    A. **Draw-A-Person Test – Karen Machover (49, 51)**

    B. **House-Tree-Person**   **(House: life, Tree: environment, Person: interpersonal relationships)**

V. **Projective Paradox**

    A. **Cling to Stereotypes**

    B. **Auxiliary to Clinical Interview**

Self-Report Personality Inventories

I.   **Theory Guided Inventories**

    A.   EPPS (Edwards-Murray Needs)   p. 379, Figure 9.1

    B.   Personality Research Form – Jackson/Murray Needs

         p. 381, Table 9.1

    C.   MBTI – Jung Types   p. 382

    D.   Jenkins Activity Survey – Type A

    E.   STAI – Spielberger

II.   **Factor Analytically Derived Inventories**

    A.   16 PF   (p. 386, Table 9.2)

    B.   Eysenck Personality Questionnaire

    C.   Comrey Personality Scales

    D.   NEO Personality Inventory – Revised   (NEO PI-R)

         p. 390, Table 9.4

III. **Criterion Keyed Inventories**

    A.   MMPI-2   p. 394, Table 9.5

    B.   CPI   p. 397, Table 9.6

    C.   MCMI-III   p. 398, Table 9.7

    D.   PIC-2   p. 400, Table 9.8

IV   **Assessment of Attitude**

    A.   an evaluative component with positive or negative response of some kind

    B.   unobservable / hypothetical

    C.   assessment through unique instruments

    D.   behavioral, covert, questionnaire methods (Likert Scale)

    E.   weak links with behavior

V   **Assessment of Moral Judgement**

    A.   Moral Judgement Scale (Kohlberg)

    B.   Defining Issues Test (objective scoring)

VI   **Assessment of Spiritual and Religious Concepts**

    A.   Sigmund Freud – The Future of an Illusion (1927)

    B.   William James

         – The Varieties of Religious Experience (1902)

    C.   Allport & Ross (1967)

        1. Religious Orientation Scales – Intrinsic / Extrinsic

    D.   Daniel Batson – Religion as Quest

    E.   Spiritual Well-Being Scale: Vertical / Horizontal Dimensions

    F.   Faith Maturity Scale

    G.   Spiritual Experience Index

Behavioral and Observational Assessment   Ch. 9B

I.   Focus on behavior rather than underlying traits and
      personality – Table 9.15, p. 420

II.   Behavior Therapy and Assessment
      - Table 9.16, p.422
      A.   Contingency Management, e.g. Token Economy
      B.   Cognitive Behavior Therapies – Table 9.17, p. 425

III. Structured Interview Schedules
      SCID - Table 9.18, p. 429

IV.   Assessment of Nonverbal Behavior
      A.   visual interaction
      B.   paralinguistics
      C.   facial expression
      D.   profile of nonverbal sensitivity

V.   Analogue Behavioral Assessment
      A.   contrived setting

Neuropsychology

I.   Definition – Relationship between brain function and behavior.

II.   Anatomy

       Cranial Nerves, Table 10.1

III.   Cerebral Lobes   P. 446, Figure 10.2

   1. Occipital Lobe - Vision

   2. Parietal Lobes – Somatosensory / Integration

   3. Temporal Lobes – Auditory, long-term memory
      modulation of biological drives

   4. Frontal lobes – Executive function/motor function

IV   Cerebral Lateralization   (P. 455, Table 10-2)

V   Neuropathology

   1. Traumatic Brain Injury

   2. Neuroplastic Disease

   3. Chronic Alcohol Abuse

   4. Alzheimer's Disease

   5. Vascular Dementia (stroke)

   6. Pseudodementia (depression)

VI   **Brain Imaging Techniques**
    1.   EEG
    2.   Cerebral Angiography
    3.   CT Scan
    4,   MRI
    5.   PET

EDUCATIONAL, VOCATIONAL, & PSYCHOLOGICAL APPRAISAL

I. VALUES

A. Rokeach Value Survey   p. ~~465~~ 444 (Table 12-~~6~~ 1)

1. Instrumental - desirable modes of conduct

2. Terminal - desirable end states of existence

B. Moral Judgment Scale

1. Stages of moral development   (Kohlberg)

   Table 12-7

C. Defining Issues Test

1. Bias against Religious conservatism

EDUCATIONAL, VOCATIONAL, & PSYCHOLOGICAL APPRAISAL

II. SPIRITUAL / RELIGIOUS CONCEPTS

A. Sigmund Freud - The Future of an Illusion (1927)

B. William James

   - The Varieties of Religious Experience (1902)

C. Allport & Ross (1967)

   1. Religious Orientation Scales

      - Intrinsic / Extrinsic

D. Daniel Batson - Religion as Quest

E. Spiritual Well-Being Scale

   - Vertical / Horizontal Dimensions

F. Faith Maturity Scale

G. Spiritual Experience Index (Table 12.11; p. 477)

Industrial and Organizational Assessment

I.   **Personnel Selection**

  Complex / Predictive Validity and Legal Concerns

  A. **Autobiographical Data**

    - biodata has substantial predictive validity

  B. **Employment Interview – poor reliability / validity**

  C. **Cognitive Ability Tests**

    - **Wonderlic Personnel Test (g)**

  D. **Personality / Temperament Tests**

    - **Hogan Personality Inventory**

  E. **Integrity Tests – Table 11.6, p. 506**

  F. **Work Sample / Situational Exercises**

    - **In-basket Test**

II. **Appraisal of Work Performance**

  A. **Performance Measures – Production**

  B. **Personnel Data: Absenteeism**

  C. **Peer Rating and Self Assessments**

  D. **Supervisor Rating Scales**

III. <u>Sources of Error in Performance Appraisal</u>

    A. **Halo Effect**

    B. **Rater Bias**

    C. **Criterion Contamination**

    D. **Guidelines – p. 520**

IV. <u>Inventories for Interest Assessment</u>

    (self-exploration and expanding career options)

    A. **Strong Interest Inventory (SII)**

      1. First published in 1927 – Strong Vocational Interest Blank. Two features:

      i) assessed likes and dislikes

      ii) empiracally keyedfor different occupations (used criterion keying)

      2. Changes in 1970's

      i) introduction of theoretical framework

      ii) renorming

      iii) increase in vocational/technical occupations

3. Current SII (317 items – only computer scored)

i) Underpinning is John Holland's work on General

Occupational Themes:

Realistic

Investigative

Artistic

Social

Enterprising

Conventional

(Person and Environment)   (RIASEC model)

ii) 25 Basic Interest Scales fall under six occupational

themes.

iii) General Occupation Scales under six occupational

themes.

iv) Problem with sample – too educated

Scoring: SS=50 SD=10

B.   Kuder General Interest Survey

1. Forced-choice format (ipsative)

2. Restricted to Adolescents

3. Ten interest areas (p.524)

C.   **Vocational Preference Inventory**

      1. **RIASEC model (Figure 11.3, p. 525)**

D.   **Self-Directed Search (SDS – Holland)**

      1. **self administered; self scored; self interpreted**

      2. **six summary scored to Realistic, Investigative, Artistic, Social, Enterprising, Conventional.**

      3. **three summary to three letter code**

      4. **compared to occupational finder**

E.   **Campbell Interest and Skill Survey   (Figure 1-5, p. 531-532)**

Forensic Application of Assessment

I. Expert Witness – well accepted measures

II. Malingering

III. Insanity Plea

    A. M' Naughten Rule, p. 540

    B. Model Penal Code

    C. Guilty but Mentally Ill

    D. R-CRAS, Table 11.8, p. 543

IV. Competency to Stand Trial - Table 11.9, p. 544

    A. Understand participants

    B. Function/cooperate

    C. Capacity to relate/understand

    D. Reasonable understanding

V. Prediction of Violence and Assessment of Risk, p. 547

    long range – short range

VI. Evaluation of Child Custody in Divorce

    A. Parent Custody Index

    B. Parent Child Relationship Inventory

VII. Personal Injury

VIII. Polygraph

Psychological Testing and the Law

I.   **Sources of Law**

    A.   Constitutional

    B.   Legislative

    C.   Judicial

II.   **Testing in School Systems**

    - Table 12.1, p.564

III. **Disability Assessment**

    A.   Public Law 94 – 142, p.569

    B.   ADA, p. 571

IV.   **Employment Testing**

    Table 12.2, p. 573

Computerized Assessment Ch. 12B

I. **Computer-Based Test Interpretation**

    A. **Scoring Report**

    B. **Descriptive Report**

    C. **Actuarial Report**

    D. **Clinical Report**

II. **Advantages/Disadvantages**

    A. **objective, fast, low cost**

    B. **risk of misuse**

III. **Computerized Adaptive Testing**

    A. **efficient measure of ability**

IV. **Future – more computerization**