Measurements

- Criterion-Referenced
  - is measuring a minimum level of acceptable performance
  - Criterion behavior

- Norm-Referenced
  - is measuring relative performance in respect with other individuals performance

RSVP

- Reliability * yields consistent results
- Standardization * is similar for all students
- Validity * actually measures what it claims to measure
- Practicality * is easy and inexpensive to use

* is the degree to which an assessment

Validity Norm-Referenced

- Content Validity
  A special case of content validity is
  - Face or Logical Validity

- Concurrent Validity
  - Correlation with criterion measurement

- Predictive Validity
  - Predictor test
  - Criterion measure

- Construct Validity
  - Trait (construct) & Performance
Reliability Norm-Referenced

- Test-Retest (practice effects)
- Parallel Forms (make two tests equal ?)
- Split-Half Method
- Kuder-Richardson  KR-21 & KR-20
  KR-21 all possible split-halves KR-21
  approximation of KR-20 is KR-21

Mean = 80, s = 10, N=100 using formula \( r_{KR} = .848 \)

Reliability Criterion-Referenced

- Cluster versus Test
- Master & Non-Master

Validity Criterion-Referenced

- Clustering (many questions) & Pre-Post
- Domain-referenced (logical validity)
- Decision Validity (Master & Non-Master)
<table>
<thead>
<tr>
<th>Factors Affecting Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Student characteristics</td>
</tr>
<tr>
<td>• Criterion measure selected</td>
</tr>
<tr>
<td>• Reliability</td>
</tr>
<tr>
<td>• Administrative procedures (clear directions)</td>
</tr>
<tr>
<td>environmental conditions</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factors Affecting Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Method of scoring</td>
</tr>
<tr>
<td>• Heterogeneity of group</td>
</tr>
<tr>
<td>• Length of test (the longer the better – fatigue?)</td>
</tr>
<tr>
<td>• Administrative procedures (clear directions)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Objectivity - Standardization</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Instructions – Trained testers</td>
</tr>
<tr>
<td>– Simple measurements – mechanical tools</td>
</tr>
<tr>
<td>– numerical values</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Administrative Feasibility - Practicality</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Cost - Time - Administration - Scoring</td>
</tr>
</tbody>
</table>
Assessment
- is the process of observing a sample of students’ behavior and make inferences about their knowledge and abilities

More assessment
- **Formative**
  - Pretests
  - CBA
    - Curriculum based assessment
- **Summative**
  - Summary of accomplishments

Using Assessment for
- Promoting Learning
  - Motivation (though extrinsic)
  - Review
  - Learning Experience
  - Feedback
- Self-Regulation
- Assessing achievement of instructional goals
• **Cognitive Domain**  (acquisition & use of information)

• **Psychomotor Domain**  (body movements and actions)

• **Affective Domain**  (attitudes and feelings)

Declarative (Content) - Procedural Knowledge

• **Behavior-Content Matrix**

<table>
<thead>
<tr>
<th>Knowledge</th>
<th>Comprehension</th>
<th>Analysis</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>History</td>
<td>Rules</td>
<td>Technique</td>
<td>Offensive Strategy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Defensive Strategy</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Understanding</th>
<th>Making Generalizations</th>
<th>Locating Information</th>
<th>Interpreting Graphs</th>
</tr>
</thead>
<tbody>
<tr>
<td>History</td>
<td>Rules</td>
<td>Technique</td>
<td>Offensive Strategy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Defensive Strategy</td>
</tr>
</tbody>
</table>

**Assessment & Grading**

• Effective Teaching & Objective Grading

• Students – Parents – Teachers – Administrators
  - educational objectives tied with grading

• Including

  **Affective Factors**
  – Sportsmanship
  – Attendance
  – Participation
  – Effort
  – Showering

  **Cognitive factors**  (strategy - rules)

  **Psychomotor factors**  (Activity – Game Performance – Fitness)
Authentic (Traditional)

- Formal Record Keeping
- Natural Settings
- Formative and Summative
- Technique (Form) and End-Result
- Self-Assessment & Peer-Assessment
  - Statistics - Charts
- Portfolio Assessment

Criteria for Grades

- Educational Objectives
- Validity – Reliability – Objectivity
- **Weighted factors**
  - understandable
- Discriminate
- Administrative economy

Grading Methods

- **Norm-Referenced**
  - Natural Breaks
  - S.D.
  - Percentage
  - Norms-Methods

- **Criterion-Referenced**
  - Contract Method
  - Percent-Correct Method

- Weighting of Factors
- Reporting Final Grades (L, %, P-F, Descriptors)
Test Construction

- First step is to ensure **Content validity**
- **Test** (Item construction - Administration - Item Analysis:
difficulty - discrimination - response quality 2-3% - Revision)

Paper-Pencil Knowledge-Tests

- Objective Items
  - Recognition
- True-False
- Multiple-Choice
- Matching Items
- Short-Answer & Completion
  - Recall
- Subjective
  - Recall & Synthesis
- Essay

Psychomotor Domain Test

**Know what is a good test first**

- Physical
- Motor
- Fitness
- Play
Next steps

- Analyze performance
- Review Literature
- Devise test items
- Prepare directions
- Review tests (peers)
- Pilot the test (small group of students)
- Determine Validity, Reliability, objectivity
- Develop norms
- Intercorrelations (items related with objective – not with each other)

Testing

- **Pretest**
  Schedule – prepare instructions – scorecards – safety

- **Test**
  give instructions – demonstrate if necessary

- **Posttest**
  Grade – percentiles & statistics – interpret results – feedback

  - Motor Ability
  - Motor Capacity and Motor Educability

Psychomotor Domain

- 15. Body Composition
- 12. Cardiorespiratory Endurance - Fitness
- 14. Muscular Endurance
- 14. Muscular Strength
- 13. Flexibility
- 17. Physical Fitness
- 19. Sports Skills
- 20. Affective Behavior
- 18. Special Populations

- 14. Power
- Speed
- Reaction Time
- 10. Agility
- 11. Balance
- Kinesthetic Perception
- Coordination
- 16. Posture
1. A test is used when individuals are expected to perform at a specific level of achievement.
2. A test is used when an individual's performance on a test is compared with other individuals' performance on the same test.
3. Refers to the degree to which a test actually measures what it claims to measure.
4. A test's validity is related to how well a test measures all skills and subject matter that have been presented to a group.
5. A test's validity is concerned with future performance.
6. A test has validity when it obviously measures the desired skill.
7. Validity indicates how well an individual performs a skill at the present time.
8. Validity refers to the degree that the individual possesses a trait presumed to be reflected in the test performance.
9. Validity is often used to estimate how well a test measures what it claims to measure.
10. The validity of a test may be affected by .
11. Refers to the consistency of a test.
12. The reliability of a test can be determined .
13. The reliability of a(n) test is defined as how consistent the test classifies individuals as masters or nonmasters.
14. The reliability of a test may be affected by .
15. A test has high when two or more persons can administer the same test to the same group and obtain approximately the same results.
16. The first step in planning knowledge test items is to .
17. All knowledge test items should .
18. An index of or above on a norm-referenced test indicates that the item discriminates well.
19. Generally, each response for multiple-choice item should be selected by at least percent of the test-takers.
20. For best results, assessment of students should be performed the teaching of a skill.
21. Authentic assessment allows the teacher to .
22. Formative assessment is performed .
23. Summative assessment is performed .
24. For assessment purposes, a portfolio may include .
25. Grades are symbols that denote .
26. Grades for physical education should .
27. Before you construct a psychomotor test, you should .
28. When you devise items for a psychomotor test, you should make the .
29. Concurrent validity for a norm-referenced test can be estimated through .
30. The first step in the construction of a rating scale is to .
31. Anthropometry may include measurement of .
32. For measurement purposes, body composition refers to .
33. A male is considered obese when his body fat content exceeds % of his total body weight.
34. A female is considered obese when her body fat content exceeds % of her total body weight.
35. A person with slender, long arms and legs and little muscle definition has characteristics.
36. An athletic-looking individual has characteristics.
37. A major drawback to all height-weight tables is that some individuals do not .
38. One of the most valid methods of determining body composition is through .
39. Body mass index (BMI) provides an indication of .
40. A female waist-to-hip ratio above is linked to greater health risks.
41. A male waist-to-hip ratio above is linked to greater health risks.
42. Proper posture .
43. Good posture depends on .
44. The term fitness includes .
45. Health-related physical fitness includes .
46. Skill-related physical fitness includes .
47. The term "special populations" is used when referring to individuals.
48. Evaluation of special populations should be performed in the school physical education program.
49. When testing special populations, tests should be used.
50. The most popular reason for sports skill measurement is to .
51. Affective behavior involves the of an individual.
52. Measurement of affective behavior can be used with groups .

http://www.csun.edu/~kv61497/K302/miller01.pdf

For chapters 10-20 also you should know the definitions of each ability and why we measure it.

AAHPERD shuttle run 12-minute run 1.5 mile run Sit-and-Reach
Sit-ups Push-ups etc Lordosis-Lumbar , Kyphosis-Cervical, Scoliosis-lateral curve,
(endo ecto meso)- morphic

Fitnesgram Test Components
1 Aerobic 1 mile Pacer 2 .Body Composition 3 ..... to 6

ACSM Fitness Test President’s Challenge Test Components
Bruiniks-Oseretsky BMAT

Lordosis-Lumbar , Kyphosis-Cervical, Scoliosis-lateral curve,
<table>
<thead>
<tr>
<th>Title: Definition</th>
<th>Why should we measure it</th>
<th>Name 3 tests</th>
<th>Describe one test</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>