Suggestions for assembling items

• Clean up errors -- proofread!
• Keep items free from racial, ethnic, and gender bias
• Include clear directions -- don’t assume students know these
• Group items according to item format

Suggestions for administering items

• Make students aware of any time limits but assure them adequate time is provided
• Make sure students know how to record answers
• Explain guessing -- typically have students answer all questions
• Don’t hold students on the starting line too long
• Avoid interruptions
• Avoid giving hints that might advantage some students
• Discourage cheating

Purpose of Grading

• Grading is often used to serve multiple purposes:
  – Administrative purpose of evaluating students and ranking
  – Educational purpose of assessing learning and progress toward class objectives
  – Some people also view grades as motivators
Shortcomings of Letter Grades:
• Are typically a combination of achievement, effort, work habits, and good behavior
• The proportion of students assigned each letter grade varies from teacher to teacher
• They do not indicate a student’s specific strengths and weaknesses in learning

Reliability of classroom measurement
• Research shows classroom measurement to be less than perfectly reliable
  – Different teachers asked to grade the same essays often come up grades that differ by several letter grades
  – Numerical grades are notoriously unreliable

Norm vs. criterion-referenced testing
– Norm-referenced testing’s primary purpose is administrative—ranking students for selection into programs, tracks, etc.
– Criterion-referenced testing is typically the most useful way to assess learning/progress toward educational goals
3 Options for Grading Systems

- Relative grading (norm-referenced) -- performance in relation to other group members
- Absolute grading (criterion-referenced/mastery) -- performance in relation to specified standards
- Improvement/ability grading -- performance in relation to some determined baseline or starting point

Absolute Grading

- A = Outstanding. Student has mastered all the course’s major and minor instructional goals
- B = Very good. Student has mastered all the course’s major instructional goals and most of the minor ones.
- Etc.

- A = 95% to 100% correct
- B = 85% to 94% correct
- C = 75% to 84% correct
- D = 65% to 74% correct
- F = below 65% correct

Assigning grades

- Numerical (%) grades are common in education, both today and historically. In order to do this sort of grading:
  - We must assume that there are REAL differences between grades (i.e., that an 82% is different than an 83%)
  - We must assume that the number represents something REAL
  - We must assume that our measurement is RELIABLE
One-chance testing

- One-chance testing is the norm throughout preK-postsecondary education. In one-chance testing the students take the test, receive grades, and move on.
  - What is the message here? Ever hear a student, when asked about material studied previously, say something like “Oh, we’re done with that”?
  - Teachers often assume that students will review what they missed and learn the material. Do they?
  - Is one-chance testing serving an educational or evaluative (administrative) purpose?

In other areas of “life” involving skill-acquisition is it one-shot evaluation?

- Driving test?
- Athletics?
- Basic (military) training?
- Learning to use a computer?
- Pottery making?
- Learning a language?
- Learning to walk?
- Learning to play a musical instrument?
- Learning to cook/bake?
- Learning a trade?

One of the most important abilities to develop in schools is metacognitive ability—the ability to accurately reflect on your learning and employing increasingly effective strategies.

Suggestions for Grading:

- Assess frequently
- Use true FEEDBACK to facilitate learning, base feedback on frequent or constant observations (homework, one-minute quiz, behavioral observation, peer mentoring, etc.—think coaching/mentoring here)
Self-Report Tests

• Keep in mind that self-report inventories assumes that individuals are WILLING AND ABLE to report accurately--this may be a big assumption!
• Try to also collect some observable measure of the trait of interest

Interest Inventories

• Strong-Campbell Interest Inventory -- suggests particular occupations based upon responses
• Self-Directed Search (SDS) Career Explorer -- classifies individual according to occupational themes (realistic, investigative, artistic, social, enterprising, conventional)

Personality

• Personality refers to a person’s unique and relatively stable pattern of thoughts, feelings, and actions
• Personality is an interaction between biology and environment
  – Genetic studies suggest heritability of personality
  – Other studies suggest learned components of personality
Measures of Personality

Personality refers to a person’s unique and relatively stable pattern of thoughts, feelings, and actions (traits)

- Interviews
  - Unstructured: “Tell me about yourself...”
  - Structured: Set list of questions
- Observation: Psychologist learns about personality by observing the person
- Objective tests: self-inventories that involve paper and pencil tests
- Projective tests: subjects reveal aspects of their personality when they talk about ambiguous stimuli

Personality Measurement Issues

- Objective self-report personality tests can be criticized on the basis of
  - Deliberate deception and social desirability bias
    - Can the test detect deception and attempts to enhance social desirability?
  - Inappropriate use: when tests are used for purposes other than their designed use
    - Use of a personality test to decide a presidential election

The “Big 5”

- Modern personality research argues for 5 basic personality traits (OCEAN)
  - Openness: whether a person is open to new experiences
  - Conscientiousness: whether a person is disciplined and responsible
  - Extroversion: whether a person is sociable, outgoing and affectionate
  - Agreeableness: whether a person is cooperative, trusting, and helpful
  - Neuroticism: whether a person is unstable and prone to insecurity
Overview of the Big “5”

- Openness
- Conscientiousness
- Extraversion
- Agreeableness
- Neuroticism

<table>
<thead>
<tr>
<th>Low Scores</th>
<th>High Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down-to-earth</td>
<td>Imaginative</td>
</tr>
<tr>
<td>Minimalist</td>
<td>Artistic</td>
</tr>
<tr>
<td>Liberal</td>
<td>Original</td>
</tr>
<tr>
<td>Unconventional</td>
<td>Creative</td>
</tr>
</tbody>
</table>

- Neuroticism: Emotional Instability
  - Calm
  - Controlled
  - Conscientious

- Openness: Creativity
  - Inquiring
  - Adventurous

- Agreeableness: Trust
  - Trustful
  - Cooperative

- Conscientiousness: Dutiful
  - Dutiful
  - Conscientious

- Extraversion: Sociable
  - Outgoing
  - Assertive

Projective Tests

- Projection is an idea developed by Freud in which people are thought to reveal their true feelings and thoughts when describing ambiguous stimuli.

- A projective test presents a series of ambiguous stimuli and asks that a subject describe each stimulus.
  - The idea is that their verbal descriptions will reveal key aspects of their personality.

Specific Projective Tests

- Rorschach test
  - Consists of 10 inkblots
  - Reliability and validity of this test is low.

- Thematic Apperception Test (TAT)
  - TAT also consists of a series of ambiguous figures
  - Was devised to measure achievement motivation by Henry Murray in 1938.
TAT

- That makes me think of the garden.
- It is the city in the country, very much so.
- It looks like New York, with the Empire State Building right there.
- Calming, relaxing. There’s a tree there so you can see the country-side and you’ve got the background with the city and the buildings, so it’s a regional focus.

Rorschach Inkblot Test