Assessment, Scoring, and Evaluation

- Pass/Fail
  - A
  - A-
  - B
  - C+
  - C
  - 85%
  - 93%
  - F
  - D
  - Unsatisfactory

- 67%
Definitions

- **Assessment** -- The process of measuring something with the purpose of assigning a numerical value.

- **Scoring** -- The procedure of assigning a numerical value to assessment task.

- **Evaluation** -- The process of determining the worth of something in relation to established benchmarks using assessment information.
Assessment Types

- **Formative** - for performance enhancement
- **Formal** - quizzes, tests, essays, lab reports, etc.
- **Traditional** - tests, quizzes, homework, lab reports, teacher
- **Summative** - for performance assessment
- **Informal** - active questioning during and at end of class
- **Alternative** - PBL’s, presentations, essays, book reviews, peers
Alternative Assessment

Alternative to what? Paper & pencil exams

Alternatives:
- lab work / research projects
- portfolios
- presentations
- research papers
- essays
- self-assessment / peer assessment
- lab practical
- classroom “clickers” or responder pads
More Formal Alternatives

- Rube Goldberg projects
- bridge building / rocketry / mousetrap cars
- writing a computer program
- research project
- term paper
- create web page
- create movie
- role playing
- building models
- academic competitions
Informal CATs (Classroom Assessment Techniques)

- Quick-fire questions
- Minute paper
  - 1) What did you learn today?
  - 2) What questions do you have?
- Directed paraphrasing (explain a concept to a particular audience)
- The “muddiest” point (What is it about the topic that remains unclear to you?)
Authentic Assessment

- The National Science Education Standards draft (1994) states, "Authentic assessment exercises require students to apply scientific information and reasoning to situations like those they will encounter in the world outside the classroom as well as situations that approximate how scientists do their work."
Assessment Concerns

- **Validity** -- Is the test assessing what’s intended?
  - Are test items based on stated objectives?
  - Are test items properly constructed?

- **Difficulty** -- Are questions too easy or too hard? (e.g., 30% to 70% of students should answer a given item correctly)

- **Discriminability** -- Are the performance on individual test items positively correlated with overall student performances? (e.g., only best students do well on most difficult questions)
Evaluation Types

- **Criterion-referenced evaluation** -- student performance is assessed against a set of predetermined standards.

- **Norm-referenced evaluation** -- student performance is assessed relative to the other students.

- **The “curve”** -- sometimes a combination of criterion- and norm-referenced processes.
Criterion-Referenced Eval’s

- Based on a predetermined set of criteria.
- For instance,
  - 90% and up = A
  - 80% to 89.99% = B
  - 70% to 79.99% = C
  - 60% to 69.99% = D
  - 59.99% and below = F
Criterion-Referenced Eval’s

**Pros:**

- Sets minimum performance expectations.
- Demonstrate what students can and cannot do in relation to important content-area standards (e.g., ILS).

**Cons:**

- Some times it’s hard to know just where to set boundary conditions.
- Lack of comparison data with other students and/or schools.
Norm-referenced Evaluation

- Based upon the assumption of a standard normal (Gaussian) distribution with $n > 30$.
- Employs the $z$ score:
  - A = top 10% ($z > +1.28$)
  - B = next 20% ($+0.53 < z < +1.28$)
  - C = central 40% ($-0.53 < z < +0.53$)
  - D = next 20% ($-1.28 < z < -0.53$)
  - F = bottom 10% ($z < -1.28$)

$$z = \frac{X - \bar{X}}{\sigma}$$
Norm-referenced Evaluation

- **Pros:**
  - Ensures a “spread” between top and bottom of the class for clear grade setting.
  - Shows student performance relative to group.
  - **Con:** In a group with great performance, some will be ensured an “F.”

- **Cons:**
  - Top and bottom performances can sometimes be very close.
  - Dispenses with absolute criteria for performance.
  - Being above average does not necessarily imply “A” performance.
Norm and Criterion Compared

- **Norm-Referenced:**
  - Ensures a competitive classroom atmosphere
  - Assumes a standard normal distribution
  - Small-group statistics a problem
  - Assumes “this” class like all others

- **Criterion-Referenced:**
  - Allows for a cooperative classroom atmosphere
  - No assumptions about form of distribution
  - Small-group statistics not a problem
  - Difficult to know just where to set criteria
The “Curve”

- The “curve” might represent a mixture of norm- and criterion-referenced grading.
- The “curve” is a highly subjective process.
- The “curve” is normally applied only at the end of a term.
Assessing Scientific Process Skills

- Scientific Process Skills (intellectual skills closely associated with inquiry learning)
Basic Scientific Process Skills

- Observing
- Communicating
- Classifying
- Measuring
- Inferring
- Predicting
Integrated Scientific Process Skills

- Identifying variables
- Constructing a table of data
- Constructing a graph
- Describing a relationship between variables
- Acquiring and processing data
- Analyzing investigations
- Constructing hypotheses
- Defining variables operationally
- Designing investigations
- Experimenting
Enhanced Scientific Process Skills

- Solving complex, real-world problems
- Establishing empirical laws
- Synthesizing theoretical explanations
- Analyzing and evaluating scientific arguments
- Constructing logical proofs
- Generating principles through the process of induction
- Generating predictions through the process of deduction
Miscellaneous Comments

- Study guides can be created to set objectives.
- Prepare tests from objectives.
- Assess broad spectrum: content AND skills.
- Make a rubric for questions that do not have forced-choice answers.
- Create an answer key for forced-choice questions.
- Double-check your answer key.
- Grade ASAP, providing corrective feedback.
Handling Appeals

- Encourage students to learn from their mistakes.
- Accept appeals in writing, due by a certain date.
- Refuse to discuss question if student will be appealing the answer.
- Appeals include the following:
  - Question being appealed
  - Teacher and student responses
  - Explanation of why student’s response is as good as or better than teacher’s expected response.
  - Teacher responds in writing.
  - No class-wide correction: each student must make own appeal.
- Benefit: students feel they are treated fairly.
**Portfolios of Student Work**

- Have students prepare an ongoing, extensive portfolio of their work.
- Maintain these portfolios in an open but supervised setting.
- During parent-teacher conferences, have student in attendance and have parents go through portfolio with student under the watchful eyes of the teacher.
Record Keeping

- Keep copies of your grade book or computer program in widely separated locations.
- Keep up to date.
- Respect confidentiality laws....