Testing & Grading
The Dark Side of Teaching or Path to Enlightenment?

Objectives
To develop a framework based on ethical and professional considerations for making testing decisions

To understand the role of institutional and local context on testing and grading decisions and select most appropriate tests for contexts

To delineate differences between criterion- and norm-based assessment and the implications of use

Why Test? Why grade?
• For the institution
  - Keep track of student progress
  - Evaluate student work for third party
  - Informing and motivating student education
  - Academic honors, awards, funding
  - Graduation

Why test? Why grade?
• For the instructor
  - Find students’ strengths and weaknesses / base knowledge
  - Find weaknesses / effectiveness in instruction
  - Accountability of instructor - evaluation
  - To help department/school/institution make decisions

Why test? Why grades?
• For the student
  - Feedback on learning - find weaknesses and areas of strength (formative)
  - Motivation to study (non-ideal)
  - Prediction of future (course) success
  - Professional preparation: licensing, certifications, etc.
  - Access to grad school, funding, awards, etc.

Preexisting Conditions:
• Students enter with different levels of knowledge, aptitude and motivation

• Faculty enter with different levels of experience in teaching, testing and grading

• Faculty receive little/no training in grading effectively, fairly and accurately

• Tests and grades have high-stakes outcomes for students
Pre-existing conditions:

• Students enter with different levels of knowledge
  Faculty must base grades on what students do in the course—only responsible for what he/she can control, i.e., present performance, not past or future (Rocheleau & Speck, p. 80, 2007)

Pre-existing conditions

• Faculty enter with different levels of experience in teaching, testing and grading
  Peer mentoring for new faculty, access to previous materials used in course, departmental standards, testing services, ITL?

Pre-existing conditions

• Faculty receive little/no training in grading effectively, fairly and accurately
  ITL resources, web-based guidelines, workshops for departments, department-based training, peer mentoring

Pre-existing conditions

• Tests and grades have high-stakes outcomes for students
  Accuracy and fairness: tests/grades that reflect mastery of material AND show performance in relation to other students.

Pre-existing conditions

• Tests and grades have high-stakes outcomes for students
  “Fairness in grading involves the idea that grades match knowledge, ability, or learning…not luck. Imprecise assignments lead not only to guesswork on students’ part, but also to inconsistent standards when grading.” (Rocheleau & Speck, p. 94, 2007)
New problems / issues

- "Despite often low levels of preparedness, students tend to be highly confident in their abilities."
- Trend toward grade inflation in high schools with fewer than 6 hours/week studying
- Millennials and entitlement
- More first generation and immigrant students

So what can you do?

Clear expectations
- syllabus
- rubrics
- assignments
- define “study” in your discipline
Clear Feedback
- formative (ungraded) and summative feedback
- more frequent testing (self-tests)
- clear grading criteria

Set the students up for success

- Test early and often. (Nilson, p. 199)
- Measure only what has been taught
- Provide point scores for items e.g. (2)
- Provide thorough review
- Provide intermediate indications of success
- Use ungraded work to give formative feedback

To promote learning:

- Return tests in timely manner and discuss results
- Allow retakes of earlier tests (even ungraded) to encourage improvement before going on with new material.
- Essays: 2 scores-content and grammar
- Include review items from previous units on each test
Note: Please make sure that all work is shown and that your answers are clear and legible. If an explanation is required or a difference is requested, please do not just define terms. Partial credit cannot be given for incorrect answers if your method is not presented clearly. Units must be included in answers for full credit. The total marks for this exam will represent about 10% of your final grade.

This is to help the students to be successful

Common misinterpretation:

Instructor: High scores=good teaching
Low scores=poor students

Students: High scores=smart students
Low scores=poor teaching

To avoid misinterpretation

• Do not interpret small differences in scores as real differences in achievement
• Final grades must be based on >one or two exams:i.e.,projects, labs, papers, proposals, etc.

Validation and evaluation of test

- Compare test results with each student’s general abilities
- Get student feedback on exam and course
- Keep records of items missed
  If missed by both high and low achievers there is probably something wrong
  If missed by mostly low achievers then it is likely valid, reliable, & discriminating
  If answered correctly by both, then it may be too easy
Assigning grades
- Base evaluation on specific skills or competencies the course is promoting.
- Define the task clearly.
- Identify criteria on which successful performance will be judged and specify criteria in advance (rubrics).
- Uniform testing and criteria for all students. (unless documented learning disabilities)

**norm-based grading** (Prichard & Sawyer)
Disadvantages:
1. No matter how outstanding the reference group of students is, *some will receive low/high grades*. Grades are difficult to interpret without additional information about the overall quality of the group.
2. Grading standards tend to fluctuate with the quality of each class of students. Often a student’s grade depends on who was in the class.

**norm-based grading**: B. Bloom
Disadvantages:
“*The normal curve is a distribution most appropriate to chance and random activity. Education is a purposeful activity and we seek to have students learn what we teach. Therefore, if we are effective, the distribution of grades will be anything but a normal curve. In fact, a normal curve is evidence of our failure to teach.*"

**criterion-based grading** (Prichard & Sawyer)
Advantages:
1. Course goals and standards must necessarily be defined clearly and communicated to the students.
2. Most students, through hard work and adequate instruction, can obtain high grades. Focus is goal achievement not competition.
3. Final grades reflect course goals.
4. Students do not jeopardize their own grades by helping other students with course work.

**criterion-based grading** (Prichard & Sawyer)
Disadvantages:
1. It is difficult and time consuming to determine what *course standards* should be for each possible course grade issued.
2. The instructor *has to decide* on reasonable expectations of students.
3. Inexperienced instructors may be at a disadvantage in making these assessments.

Some help!
--Grade inflation: new website: uwyo.edu/ctl/gradeinflation/index.asp
--Multiple Choice questions that test deep understanding: http://cte.ku.edu/teachingInnovations/gallery/visibleknowledge/eggleston/index.shtml
--How To Prepare Better Tests: Guidelines for University Faculty (handout/pdf file)
References: