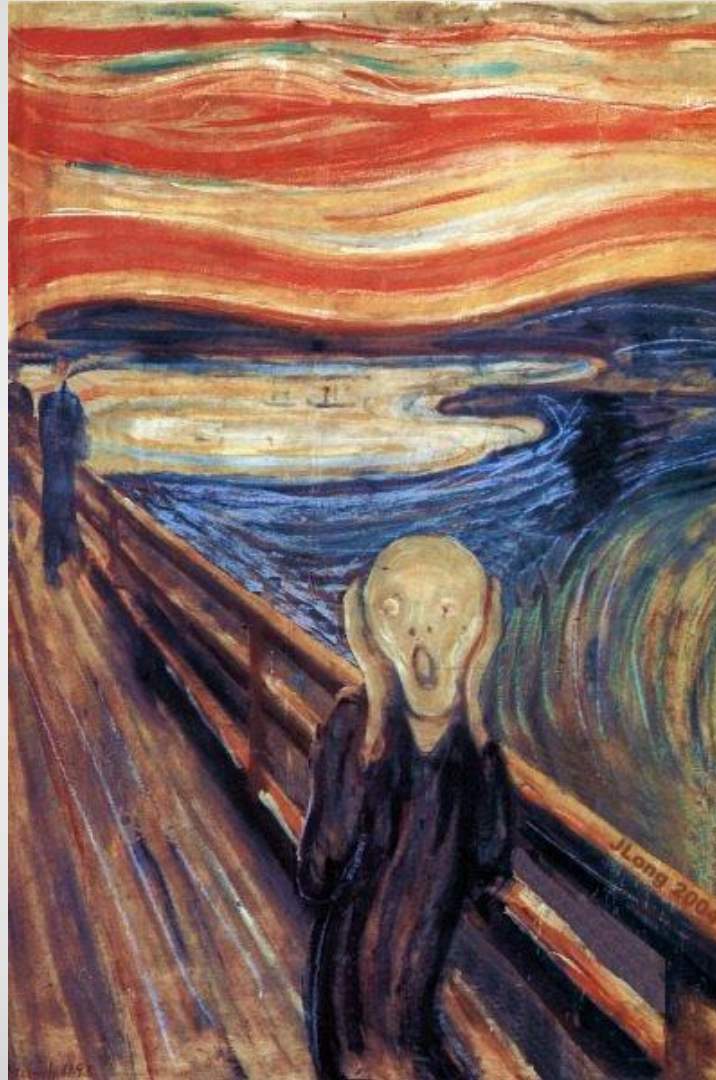


# **Assessing Student Performance: Hints for Survival**

David Morse  
Professor and Interim Head  
Counseling and Educational Psychology

# It's Time to Assess: Now What?



**Be careful what you  
ask...**

# The Test We've All Taken

1. Tom sold 140 bunches of bananas. He put 4 bananas in each bunch. How many bunches did he sell?
2. Jasmine bought six CDs for \$13.99 each. How much money did she have to spend?

(Fill in the blank to complete each statement)

3. Concave mirrors can \_\_\_\_\_ images.
4. There have been great advances in the field of \_\_\_\_\_.
5. The U.S. has a \_\_\_\_\_ economy.
6. All Mexicans are \_\_\_\_\_.
7. Describe the career of Hitler. (Essay item)

*Match the following:*

\_\_\_\_\_ 8. *Au*

a. Number of colors in the rainbow.

\_\_\_\_\_ 9. 8

b. Doctor who invented oral polio vaccine.

\_\_\_\_\_ 10. Dr. Sabin

c. Symbol for gold.

...What went wrong here?

# The assessment playing field

- Deciding *what* to assess
- Deciding *how* to assess
- Test construction
- Performance assessment
- Grading and marking
- Feedback
- Cheating, appeals, attendance, reporting

# Deciding What to Assess

- What are the *important* outcomes of the course?
  - What skills will Ss need in later in this course? In subsequent courses?
  - Are some skills more frequently used?
  - Remember, *balance*: The right location between “Everything in the book” and Guido Sarducci’s 5-minute university.

## ...what to assess

- Inform Ss of important course objectives
  - Research uniformly shows that Ss prefer this, and performance is often enhanced.
  - How?
    - Syllabus
    - Study guides
    - Sample tests or practice quizzes
    - Posted test “banks”



# Deciding How to Assess

- Issues that may affect your choices:
  - Task authenticity
  - Time required (preparation, administration, evaluation)
  - Class size
  - Frequency (more is better)
  - Formative vs. summative assessment
  - Can Ss self-appraise? (Kruger-Dunning)

## ...how to assess

- Open book vs. closed book tests  
Open book tests more likely to:
  - reduce anxiety
  - reduce need to memorize information
  - measure “higher order” skills (if so built)
- Take-home tests:  
Ss do not learn or retain as well as when in-class tests are used.
- On-line tests  
Security, suitability, quality
- Collaborative tests

# Some Testing Issues

- Frequency of testing
  - Frequent testing appears to be beneficial, especially for lower ability Ss
  - Many Ss appear to prefer frequent testing schedule
- “Pop” tests: generally not recommended
  - Ss perform better on announced tests.
- Order of difficulty: mixed findings.
  - Many studies suggest no difference; most recommend easy to hard sequence

# Test Construction

- Use publisher's "test banks" with caution--they are often poorly done.
- Show your tests to someone else before using. *Be prepared to scrap/revise items.*
- A good test is one that:
  - 1. Measures the intended behavior(s)
  - 2. Clearly communicates task(s) to Ss
  - 3. Yields reliable indication of performance.



# Steps in Test Construction

1. Identify primary purpose(s) for test scores.
2. Identify behaviors that represent construct or define domain.
3. Prepare test specifications; item specifications.
4. Construct initial item pool.
5. Have items reviewed; revise as necessary.
6. Preliminary item tryouts; revise as necessary.
7. Large-scale field test (with representative sample).
8. Statistical analysis of items, scores; eliminate poor items.
9. Develop guidelines for administration, scoring, interpretation of scores.

- **Constructed response**
  - Short answer, completion, essay, performance tasks.
  - Often easier to construct
  - Slower response
  - Slower scoring
  - Good for “higher-order” skills
- **Selected response**
  - True-false, multiple choice, matching
  - Quicker response
  - Quicker, more objective scoring
  - More challenging to measure “higher-order” cognitive skills

## Test Format Implications

# General Rules for Item Writing

- Have others review your work.
- Be prepared to revise items.
- Include clear directions for examinees.
- Avoid questions from direct quotations.
- Avoid trick questions.
- Have items reflect importance, time spent.
- Avoid double negatives

# Avoid "Give-away" Questions

(note: keyed answer indicated by asterisk)

1. The primary purpose of the class in frumpaling is to
  - a. remove class-prangs. (\*)
  - b. patch tremalls.
  - c. loosen cloughs.
  - d. repair plumots.
  
2. The fribbled breg will snicker best with an
  - a. Mors.
  - b. Ignu. (\*)
  - c. Derst.
  - d. Sortar.



# ...more examples:

(note: keyed response indicated by asterisk)

1. Mitushito was:
  - a. a former Japanese Emperor. (\*)
  - b. a U.S. delegate to the United Nations.
  - c. an English pilot.
  - d. Russian journalist.
  
2. Colchicine treatment of Jimson weed may result in twice the number of chromosomes, thus being a cause of change in the offspring. This type of change is called a(n):
  - a. mutation.
  - b. polyploid. (\*)
  - c. segregation.
  - d. hypotrichosis.

# Avoid Debatable Questions\*

Who was the most important governor of Mississippi?

What is the best thing about America's economy?

Who was the better poet, Wordsworth or Keats?

Was Stanley Milgram's research on *S* obedience immoral?

\*(Unless your intention is to measure debating skill.)

# Avoid Vague Tasks or Questions

List the results of World War II. Use both sides of the paper if necessary.

Explain the symbolism in *Moby Dick*.

Explain the language of society. (Sociology course)

*Don't use essays for simple listing tasks, e.g.:*

List the confederate states and give the dates of secession.

# Avoid Trick Questions

T F John L. Tukey developed the display known as the box and whisker plot.

T F Abraham Lincoln, the 16th President, abolished slavery in the states in the Emancipation Proclamation.

T F J. Presper Eckert is one of two persons credited with inventing the modern electronic computer.

# Be sure tasks will indicate skill attainment

*What's wrong with this item? Capital is Pierre.*

The capital of South Dakota is:

- a. Bismarck
- b. Cheyenne
- c. Boise
- d. Montpelier
- e. None of the above

# Evaluating Ss Work/Answers

- Remove names from your sight.
- Prepare model answer / scoring guide in advance.  
Partial credit possible?
- After completion, re-read some early papers to assure consistent standards.
- Take breaks.
- Read same question for all; move on.
- Use multiple raters (if possible).

**Be Careful How You  
Ask...**

# Why It Helps to Prepare Questions and Keys in Advance...

Don't say I didn't warn you!





**From physics...**

- Tie string to barometer, stand on top of building, lower barometer to ground. Measure string; add to barometer height.
- Stand on top of building, drop barometer and time the fall. Use acceleration of gravity to solve for height.
- Climb stairs, marking off building height in “barometer units”
- Measure shadow of barometer while on ground; measure shadow of building. Solve for height by proportions.
- Offer custodian barometer if he tells you building height.

Where was the American Declaration of Independence signed?

*At the bottom.*

**...you know what I meant!**

## Transparency Worksheet 23 Hard and Soft Water

1. Briefly explain what hard water is.

ice

2. Note that calcium is one of the solids dissolved in ocean water. Describe two ways by which calcium and up is sedimented out

**...and your problem is?**

1.21

4c) Expand

~~$x^2 + x - 2$~~

$(a+b)^n$

*Very young Peter.*

$= (a + b)^n$

2 ?

$= (a + b)^n$

$= (a + b)^n$

~~X~~

etc...

2. A 3-kg object is released from rest at a height of 5m on a curved frictionless ramp. At the foot of the ramp is a spring of force constant  $k = 100 \text{ N/m}$ . The object slides down the ramp and into the spring, compressing it a distance  $x$  before coming to rest.

10 (a) Find  $x$ .

5 (b) Does the object continue to move after it comes to rest? If yes, how high will it go up the slope before it comes to rest?



$$U = 3(9.81)(5) = 147.15$$

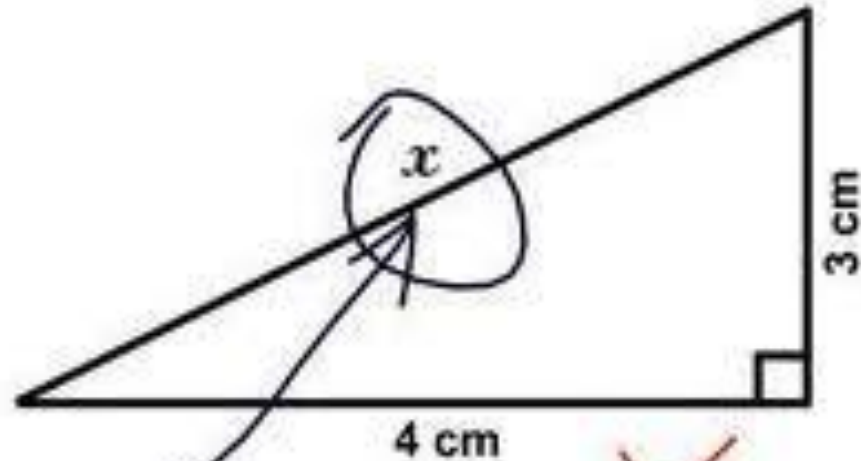
$$U_s = \frac{1}{2}(100)x^2 = 50x^2 \quad \dots?$$

NO. there is an elephant in the way.





3. Find  $x$ .



Here it is



Ocular Trauma - by Wade Clarke ©2005

...I knew that!

Charles Darwin was a naturalist who wrote the organ of the species.

Benjamin Franklin produced electricity by rubbing cats backwards.

The theory of evolution was greatly objected to because it made man think.

Three kinds of blood vessels are arteries, vanes and caterpillars.

The dodo is a bird that is almost decent by now.

The process of turning steam back into water again is called conversation.

When you smell an odorless gas, it is probably carbon monoxide.

A triangle which has an angle of 135 degrees is called an obscene triangle.

*...this is why a scoring guide should be prepared in advance.*



# Performance Assessment/Portfolio

Portfolio: A purposeful collection of Ss work that exhibits efforts, progress, or achievement in given area(s).

- Used often in writing, language & fine arts
- Issues:
  - Reliability of assessments may be low
  - Cost and time to evaluate or score
  - Can tap wider variety of skills than traditional measures

# Performance: Process vs. Product

- **Process:** Observations are focused on the *procedures* followed (e.g., right steps, correct sequence?)
- **Product:** Observations are focused on the measurement of the *end result* (e.g., well done, all the desired features present?)

For either, can use:

- Checklist (if behavior can be judged analytically)
- Numeric rating scale (set “anchors”)

For both, avoid ambiguity in facet(s) to be judged.

# 5 questions: Your grading philosophy

(See resources slide for Frisbie & Waltman reference)

- What meaning should each grade symbol carry? *What should “failure” mean?*
- What elements of performance should be incorporated in a course grade? *How should they be combined?*
- How should the grades in a class be distributed? *Similarly for honors, regular, remedial, lower level, upper level classes?*
- What method should be used to assign grades?
- Should borderline cases be reviewed?

# Grading and Marking

- Grades should represent performance / attainment only.
- Absolute vs. relative assessment
- To “curve” or not to curve?
  - Students’ prayer: Give us easier tests to make us look smarter, or dumber classmates!
- Combining scores\*\*

- 9 Course modules of content
- 53 Quizzes (~120 total weight: 10%)
- 9 Tests (~120 total weight: 45%)
- 1 Final exam (25-item, MC, weight: 45%)

**What do you think?**

# Feedback

- Plan assessment and feedback prior to drop date; absolutely by midterm time.
- Prompt feedback on performance is welcome and helpful.
- Give constructive feedback whenever possible, not just a score or grade.
- Do not post grades in public!



**Cheating....**

# Prevalence of Cheating

Watson & Sottile (2010) surveyed Ss at Marshall U. and reported:

No overall difference in admitted cheating between live, online classes (about 1/3)

Ss more than twice as likely to be caught cheating in live class (5% vs. 2%)

More admit receiving test help during test in distance courses (23% vs. 18%)

Intentional plagiarism more common in live classes (by 2:1 or more)

More women admit cheating in online classes (38% vs. 21%)



# Rationalizations for Cheating

- Easy to do
- No time for study
- Friend needed help
- I have to pass this course
- Everyone else does it
- No one cares if I cheat
- Someone stole my work
- Course too hard, or teacher unfair
- Course is useless

Source: B. Christe (2003)

# Tactics to Reduce Cheating

- Build good relationships with students
- Clearly define inappropriate behavior
- Discuss institutional policies
- Give reasons to play honestly
- Explain relevance of course
- Indicate how course tools can monitor compliance
- Are Ss ready (skill set; technology access)?
- Anticipate common issues (MyCourses was down)

Source: B. Christe (2003)

# Being Proactive

- Monitor student activity
- Vary types of assessments used
- Use tasks that don't lend themselves to simple "cut and paste" or Google search answers/solutions
- Assume Ss have access to other resources (*Respondus* isn't the answer)
- Time limits; large item pools; monitor
- Plan for the worst: indicate penalties; act.

Source: B. Christe (2003)

# Appeals Happen

- Make sure your syllabus is clear on how grades are formulated; *don't deviate from your published policies.*
- Save your records: At least 1 semester.
- Keep copies of tests, papers if feasible.
- Do offer to meet if a question about grade arises. Document your discussion(s). Know MSU policy.

# Don't forget these in syllabus!

- Assignments and due dates
- Late assignment policy
- Test dates; material covered
- Make-up test policy
- Cheating, plagiarism, honor code policy
- Attendance policy
- How course grades are generated.

## Examples of syllabi: Can you determine which (or which parts) are good?

- <http://faculty.fuqua.duke.edu/~mela/marketing460/grading.htm>
- <http://wings.buffalo.edu/smbbs/ana/medical.htm>
- [http://www2.arch.uiuc.edu/kanthony/arch424FA04/Syllabus\\_Details/sylGrading.html](http://www2.arch.uiuc.edu/kanthony/arch424FA04/Syllabus_Details/sylGrading.html)
- <http://www.skidmore.edu/~mmarx/idsyllaweb.html>
- <http://privacy.cs.cmu.edu/courses/pad1/syllabus.html#grading>
- <http://faculty.washington.edu/agarcia3/phys122/syllabuss.htm>

# Syllabus example on grading

From a course in Spanish (*URL omitted, to protect the guilty*):

GRADING: Grading will be based on a total points system. The final grade will be the total earned points divided by the total possible points. Tests will be graded out of 100 points and quizzes will be worth 10-50 points. Homework will be graded out of 10 possible points. Homework will be given and checked daily as preparation is necessary for learning a new language. *I expect homework to be done, and if not your final grade will be affected.*

# Attendance

- Attendance reporting required by MSU for undergraduates, Ss receiving certain financial support. Keep a record!
- Can attendance be part of course grade? If you must, try these ideas:
  - Limit to small fraction of course points.
  - Try using as incentive, not punishment.
  - Be clear about what comprises attendance.



- <http://www.webster.edu/~corbetre/haiti/voodoo/voodoo-syl-2001.htm>
- <http://www.english.upenn.edu/~jenglish/Courses/engl3syl.html>
- How about this one (from a syllabus in astronomy)?  
*“According to University policy, all students are to be present at the Final Exam time. If a student is not in attendance during the Final Exam a grade of F is assigned for the entire course, regardless of the work done by the student during the semester.”*

**Examples of attendance policies:  
Which ones would you endorse?**

# Reporting

- Mid-term grades are required for undergraduates. Post on Banner system.
- Final grades ordinarily due within a few days of final exam date. Don't be late!
- Do not openly post student grades. On-line posting is virtually immediate.

# Resources

- Faculty, former TAs. (\*)
- Instructor or publisher web sites (\*)
- Text publishers/representatives (\*)  
Review copies, instructor resources, web sites
- Measurement and evaluation texts
- See the last two slides for specific URLs  
*(\*) Take with large grain of salt.*

Now I lay me down to study,  
I pray the Lord I won't go nutty.  
If I should fail to learn this junk,  
I pray the Lord I will not flunk.  
But if I do, don't shed a tear,  
Just put a rose behind my ear.  
Tell my teacher I did my best.  
Then pile my books upon my chest.  
If I should die before I wake,  
That's one less test I'll have to take.

--Suffering Student, Cited in Ann Landers (1981, May 19).

## The Test Prayer

# Other resources on grading / marking (*from whence some of these ideas came*)

Article by Frisbie, D.A. & Waltman, K. K. (1992, Fall). An NCME Instructional Module on: Developing a Personal Grading Plan. *Educational Measurement: Issues and Practice*. See first URL:

- <http://depts.washington.edu/grading/plan/frisbie.html>
- <http://depts.washington.edu/grading/plan/index.html>
- <http://www1.umn.edu/ohr/teachlearn/resources/guides/grading/index.html>
- [http://www.egr.msu.edu/~wolff/teaching/case\\_history\\_1.htm](http://www.egr.msu.edu/~wolff/teaching/case_history_1.htm)
- <http://www.evergreen.edu/washcenter/resources/acl/index.html>
- <http://itc.boisestate.edu/tltr/disc02.htm>
- <http://www.ust.hk/celt/ta/taguide/skills/dishonesty.htm>
- <http://www.brookes.ac.uk/services/ocsd/firstwords/fw22.html>
- [http://www.utdallas.edu/dept/graddean/ta\\_handbook.htm#efficency](http://www.utdallas.edu/dept/graddean/ta_handbook.htm#efficency)

# Resources on test construction

- <http://www.delweg.com/dpwessay/tests.htm> (brief suggestions)
- <http://712educators.about.com/cs/assessment/a/assessments.htm> (brief suggestions)
- <http://www.cte.cornell.edu/campus/teach/faculty/Materials/TestConstructionManual.pdf> (very detailed, somewhat like a text chapter)
- <http://www.utexas.edu/academic/cte/sourcebook/tests.pdf> (not quite as elaborate as previous link)
- <http://pareonline.net/genpare.asp?wh=1&abt=Test+Construction> (excellent compendium of related topics)

"This class was a religious experience for me... I had to take it all on faith."

"Text makes a satisfying 'thud' when dropped on the floor."

"The class is worthwhile because I need it for the degree."

"His blackboard technique puts Rembrandt to shame."

"Textbook is confusing... Someone with a knowledge of English should proofread it."

"Have you ever fell asleep in class and awoke in another? That's the way I felt all term."

"In class I learn I can fudge answers and get away with it."

## **Course Evaluations (MIT, 1991)** **(Listen to what students are saying)**

# Cheating References

Christe, B. (2003). Designing online courses to discourage dishonesty: Incorporate a multilayered approach to promote honest student learning. *Educause Quarterly*, 26(4), 54-58.  
<http://net.educause.edu/ir/library/pdf/eqm0348.pdf>

Watson, G., & Sottile, J. (2010). Cheating in the digital age: Do students cheat more in online courses? *Online Journal of Distance Learning Administration*, 13(1).  
<http://www.westga.edu/~distance/ojdla>





**Have a great semester!**