

Better Testing = More Learning

Elena Berman, Ph.D.
Assessment and Faculty Development Specialist
Assessment and Enrollment Research
626-4214 ♦ eberman@u.arizona.edu

Evaluating Student Learning

- ♦ value of evaluation
- ♦ relating evaluation to desired outcomes
- ♦ ensuring validity and reliability
- ♦ maximizing student success

Value of Testing

"Minute for minute, testing is
the best of all teaching techniques."

LB Mirrielees, quoted in *The
Measurement of Classroom Learning*

Why is testing so valuable?

decades of research
show that students
learn what they are
tested on!

Teach to the Tests!

- ♦ *If it's YOUR testing plan,
and it's a good testing plan,*
- ♦ *then teaching to it is appropriate
and students will appreciate
your doing so*

Key Questions for Effective Evaluation of Learning

- ♦ What do you want students to
know/think/be able to do?
- ♦ How will you know if they
know/think/can do it?

Desired Learning Outcomes

- ◆ cognitive
- ◆ attitudinal
- ◆ skill

Some Commonly Used Measures

- ◆ pre/post tests
- ◆ projects, papers
- ◆ standardized tests (national, local)
- ◆ oral exams, interviews
- ◆ questionnaires, surveys
- ◆ performance measures
- ◆ portfolios

Key concepts of evaluation

- ◆ *Validity:*
you're measuring what you intend to measure
- ◆ *Reliability:*
the measures work consistently across circumstances

Ensuring Validity and Reliability

- ◆ Be sure you're testing for what you want students to know/do
- ◆ Base your course grade on multiple elements
- ◆ Use a variety of assessment measures
- ◆ Measure each major course objective using more than one method
- ◆ Provide explicit, objective criteria for grading
- ◆ Analyze test results

Sample Table of Specifications for a course

Desired Outcome	Evaluation Instrument				
	Exam 1	Exam 2	Final	HW	Project
knowledge of concept A	✓		✓	✓	
given X, will be able to do Y		✓	✓	✓	
understand rel. bet. X and Y				✓	✓
familiar with relevant literature				✓	✓
etc.					

Sample Table of Specifications for an exam

Desired Outcome	Test Questions
knowledge of concept A	1-5, 21-25
knowledge of concept B	6-15, Essay 1
given X, will be able to do Y	Essays 2,4
understand rel. bet. X and Y	Essay 3
familiar with relevant literature	16-20

Analyze Test Results

"Objective" Tests

- ◆ > 70% correct: too easy
- ◆ < 30% correct: too hard
- ◆ common errors
- ◆ faulty items
- ◆ overall test reliability

Essay Tests

- ◆ check performance on specific criteria

Helping Students Succeed

- ◆ State your grading policy in the syllabus
- ◆ Prepare students for tests/assignments
- ◆ Give the first test early in the semester
- ◆ Debrief tests/assignments with the class
- ◆ Solicit feedback about your grading system

Strategies for Enhancing the Effect of Testing on Learning*

- ◆ test material soon after it is introduced
- ◆ give frequent tests
- ◆ increase interval between tests
- ◆ make some tests cumulative
- ◆ provide immediate feedback

*F.N. Dempster

NOTES

- Strive for: Validity Reliability Reality (= "Authentic testing")
- Be sure you are testing LEARNING, not short-term memory
- Tests shouldn't be a surprise -- students should be prepared for tests; have practice with types of questions they will see, etc.

Advantages of MC Tests

- ◆ can cover a lot, yet be completed in a short time
- ◆ can test complex knowledge
- ◆ easy and reliable grading (machine-scorable)
- ◆ provide clues about student misconceptions
- ◆ can be banked for re-use; mixed and matched to produce equivalent versions
- ◆ difficulty level can easily be controlled

Disadvantages of MC Tests

- ◆ cannot test ability to organize and express ideas
- ◆ hard to give credit for partially correct answers
- ◆ subject to “clueing”
- ◆ hard to develop good items
- ◆ high probability of poorly written/poorly read items

Advantages of Extended Writing Tests and Assignments

- ◆ excellent for testing complex knowledge
- ◆ can test ability to organize and express ideas
- ◆ relatively easy to write good questions
- ◆ give scope for student individuality and creativity

Disadvantages of Extended Writing Tests and Assignments

- ◆ time-consuming to grade
- ◆ hard to grade consistently (reliably)
- ◆ hard to cover a range of material in a limited time
- ◆ penalize students with poor writing skills
- ◆ penalize non-native speakers

Advantages of Online Tests and Assignments

Disadvantages of Online Tests and Assignments

Using Pretests

Pretesting is a useful strategy for beginning a class, especially when classes contain students with mixed abilities coming from a diversity of backgrounds. Pretesting is a way of sampling where the students are relative to the content you're going to teach. It is ungraded and may take many forms, including map-making, providing samples of poetic forms, speaking in public, and responding to a case study, as well as conventional multiple choice or free response formats.

If you present pretesting as an information-gathering, or needs assessment, activity, that will not affect their grades in any way, most students will engage willingly. This is especially likely if you share with them some of the many benefits of pretesting.

Benefits of pretesting

- provides a preview of course coverage for students
- can act as an "advance organizer," which may improve student learning
- allows for early adjustment based on student interests and prior knowledge
- allows for early identification of student strengths and weaknesses
- improves ability to measure what has actually been learned in a course
- provides a way to get students started as active thinkers during classtime
- offers a potentially powerful tool for introducing students to working in small groups
- provides the instructor a good overview of class composition
- provides data that may be used in tracking student success and other research
- designing a comprehensive pretest helps an instructor review his or her vision for the course (which may or may not be the vision of the students)

Present the pretest as an information-gathering tool and use it to initiate a tradition of active thinking during classtime. Have students first take the pretest individually. Then ask them to compare answers in small groups and compile a group response. (Collect the individual responses for use in identifying advanced or deficient students.) If you plan to have students work together in groups during the semester, this is a powerful way to introduce them to the practice in a non-threatening context where the superiority of the group product will be clearly demonstrated.

Designing a good pretest requires you to have a conceptual overview of the course. This means bringing to consciousness the prerequisites and instructional goals, which makes fine-tuning easier in response to realities of student background and interests as well as exigencies of curriculum committees and departments.

Pretest data can be used in researching relations between preparedness and achievement. Data can also be used to design remedial modules, support arguments for curriculum change, and document changes in student ability that are a clear result of conscious teaching choices (prima facie evidence of teaching excellence).

What do pretests cover:

A comprehensive pretest seeks information about all three of the following areas:

1) knowledge of the course content

A pretest usually samples the important concepts of the entire course in such a way that the responses can be compared to responses to assignments or course exams. This gives the students a

preview of what they will be expected to learn, while giving you information about how much students already know. If most students are familiar with certain domains of the course content, you can address that content at a higher level. You can also use analogies from areas of greater common knowledge to explain new material. Some research suggests that previewing course content in this fashion actually increases learning by acting as an "advance organizer" for students.

It's a good strategy to collect and retain the pretests and use some of the same questions during the semester and on the final. It's often morale-building for students to see clearly that despite continuing confusion, they actually have learned quite a lot.

As students work the pretest, you can diffuse anxiety by reminding them that if they could answer all the questions, there'd be no point in taking the course.

2) attitudes about the course content, the course in general, and student goals

Questions about student attitudes, beliefs, and preconceptions provide invaluable information about areas of high/low curiosity, prejudice, prior (mis)information, etc. that you can draw on to tailor lectures and trigger discussions. An overall course strategy may need shifting if students turn out to have little interest in the course content as you've conceived it, whether the shift is to be a change of focus or a head-on confrontation with student lack of interest.

Information about student goals and interests may suggest improvements in course or assignment design. For example, a class in which student interests cover a wide spectrum may call for an assignment strategy that maximizes options, while a narrower range of interests may suggest group assignments that probe more deeply into a single subject.

3) knowledge of prerequisite material

In addition to covering the course content and student attitudes, it is important to ensure that students are able to do the work required for the course. If students lack prerequisite skills or knowledge, the sooner they find out about it, the better their chances for constructive action (whether that means dropping the course or doing remedial work). If a significant percentage lack prerequisite skills, instructors should consider changing the course to incorporate areas of deficiency.

Inform students unable to meet prerequisites that being successful in the course will be an extra challenge for them. If possible, direct them toward resources for remedying deficiencies. You will be doing both yourself and your students a favor if a deadline date is set when students are again tested and required to demonstrate basic proficiency with prerequisite knowledge.

Studies have shown that students able to solve problems or state generalizations in one realm are often baffled when asked to perform the same tasks in different contexts. That is, even if students can define common statistical terms or physical laws in general terms, they may not recognize them in contexts proper to your course. Therefore, pretesting for background knowledge should involve questions using the terminology and content of the course being taught. If you find out that knowledge transfer is a problem, you could provide a review of how terms, concepts, formulae, etc. will be used in the class.

For more information

Angelo, T.A. and K.P. Cross, *Classroom Assessment Strategies*, San Francisco: Jossey-Bass Publishers, Second Edition 1993. A compendium of useful ideas for getting feedback from students.

Glenda Wilkes, *Thinking About College Teaching*, Volume 2, Nos. 5&6, available from the University Teaching Center, 621-7788. One page papers on related topics.

**Written by Elena Berman, Ph.D., Assessment and Enrollment Research, eberman@u.arizona.edu.*

Sample Writing Assignment and Grading Rubric

The writing assignment below was used in a junior-level course on teaching and learning. Higher order thinking skills are needed to decide which materials to include given the specified purpose and audience. Note that the grading rubric, next page, lists appropriateness to audience second to content in specifying adequacy standards.

REPORT AND PRESENTATION ON NON-TRADITIONAL STUDENT GROUPS

The demographics of the university student population have changed dramatically over the past thirty years. Nationally, the traditional white, straight-out-of-high-school, middle-to-upper-middle-class, full-time student now accounts for less than 20% of the university student population nationwide. Adaptation to this changed student body has generally been ad hoc and outside the traditional academic structure.

As a class, we will undertake a major survey of diversity issues in higher education, with the goal of developing a presentation for faculty and other campus groups.

After a discussion in class, each student will choose a non-traditional student group and explore the questions below:

1. In what time periods, and in what numbers, has the group you're studying entered the university? Provide a brief history relating to the university and the group you're studying.
2. What problems or issues has your group encountered in gaining access or being successful?
3. What programs have been successful in alleviating these problems?
4. What issues or problems arise related to classroom instruction methods or faculty behavior in the classroom?
5. What changes in classroom behavior could faculty make that would improve learning for the group you're studying?

Criteria for evaluation of papers on non-traditional student populations

*Goal: an informative, readable summary of information
about a non-traditional student population, aimed at UofA faculty*

	Excellent	Satisfactory	Needs improvement
Content	<ul style="list-style-type: none"> •answers all questions fully or explains why a complete answer can't be given. •provides sources for all claims. •provides multiple examples of successful approaches that can be applied in the classroom. 	<ul style="list-style-type: none"> •provides reasonably complete responses to all questions, sources for most claims, and two or three examples of successful approaches that can be applied in the classroom. 	<ul style="list-style-type: none"> •questions are incompletely addressed; no sources provided for claims.
Appropriateness to audience	<ul style="list-style-type: none"> •information and examples are related to everyday experiences of faculty. •writing is concise; summary statements introduce sections. •awareness of audiences is demonstrated by highlighting of especially important information. •tone is positive; assumes good will and interest on the part of readers. 	<ul style="list-style-type: none"> •information is appropriate to audience and reasonably accessible. 	<ul style="list-style-type: none"> •paper is not organized for easy access to information. •paper has faculty-bashing tone.
Expression and format	<ul style="list-style-type: none"> •paper has effective sentence and paragraph structure. •paper looks attractive and approachable; format and headings direct readers to information. 	<ul style="list-style-type: none"> •most sentences are concise and show good word choice and arrangement. •most paragraphs are well-organized and appropriately restricted. •format provides reasonable access to information. 	<ul style="list-style-type: none"> •paper is excessively wordy with many poorly structured sentences and poorly organized paragraphs. •paper looks sloppy or crowded, uninviting to potential readers.
Mechanics	<ul style="list-style-type: none"> •no more than three spelling, punctuation, or grammatical errors. (This includes typos, so be sure to spell-check and then to edit.) 	<ul style="list-style-type: none"> •no more than six spelling, punctuation, or grammar errors. 	<ul style="list-style-type: none"> •more than six mechanical errors.

A Short Bibliography on Grading

Anderson, R.S. and B. W. Speck, eds. *Changing the Way We Grade Student Performance: Classroom Assessment and the New Learning Paradigm*. New Directions for Teaching and Learning No. 74, Summer 1998, San Francisco: Jossey-Bass.

Brookhart, Susan M. *The Art and Science of Classroom Assessment: The Missing Part of Pedagogy*. ASHE-ERIC Higher Education Report (Vol. 27, No. 1. Washington, DC: The George Washington University Graduate School of Education and Human Development, 1999.

Ebel, R.L. *Essentials of Educational Measurement*. Engelwood Cliffs, NJ: Prentice-Hall, 1991.

Eison, J.A., Milton, O., and Pollio, H.R. *Making Sense of College Grades*. San Francisco, CA: Jossey-Bass, 1986.

Jacobs, L.C. and C. I. Chase. *Developing and Using Tests Effectively: A Guide for Faculty*. San Francisco, CA: Jossey-Bass, 1992.

Janzow F. and J. Eison, "Grades: Their Influence on Students and Faculty." From M.D.Svinick, ed. *The Changing face of College Teaching*, New Directions for Teaching and Learning No. 40, Summer 1990, San Francisco: Jossey-Bass.

Ory, John and Katherine Ryan. *Tips for Improving Testing and Grading*, Sage Publications, 1993.

Pollio, H.R. and W. L. Humphreys, "Grading Students." From R.E. Young and K. Eble, eds., *Assessing Students' Learning*, New Directions for Teaching and Learning No. 34, Summer 1988, San Francisco: Jossey-Bass.

Tobias, S. and J. Raphael. *The Hidden Curriculum: Faculty-Made Tests in Science*, Parts 1 (Lower-Division courses) and 2 (Upper-Division Courses). New York: Plenum Press, 1997.

Walvoord, B.E. and V. J. Anderson. *Effective Grading*, San Francisco, CA: Jossey-Bass, 1998.

Wergin, J. F. "Basic Issues and Principles in Classroom Assessment." From R.E. Young and K. Eble, eds., *Assessing Students' Learning*, New Directions for Teaching and Learning No. 34, Summer 1988, San Francisco: Jossey-Bass.