



The Agony and the Equity: Testing and Grading

Why do we do it?

Tests aren't fun to take and certainly not fun to grade, and sometimes it seems life would be a lot simpler if learning—and not tests and grades—were more important.

We know, even if our students don't always agree, that exams have their purposes beyond causing panic and "all nighters." Tests help motivate students and reinforce their learning while at the same time allowing us to assess their mastery of content. They also provide us with feedback on our teaching, often showing us what was and wasn't communicated clearly. Here we'll explore some of the important learning issues involved in testing, and offer some advice on test construction and grading practices.

Base Exams on Course Objectives and Learning Goals

Students' expectations of the type of exam you'll give greatly influence how they study. If students expect a conceptually demanding exam (usually associated more with essay than with multiple choice exams, for example) then they study more deeply and their learning endures longer. Thus tests need to do more than cover content,

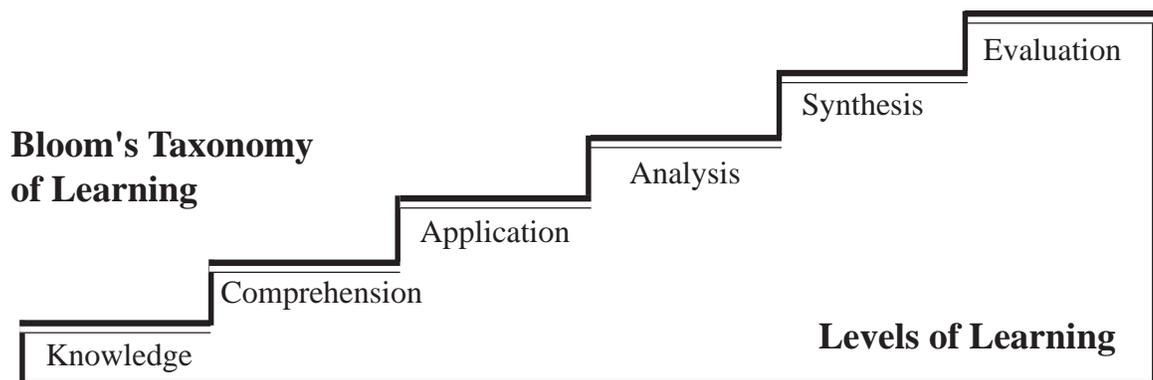
though of course they need to do this too, fairly and fully. Tests also need to cover our larger objectives for student learning—that is, what we expect our students to be able to do when they leave our classes.

In order to make sure that our tests reflect what we think our students should be learning, it's important for us to explicitly identify our course objectives. If we want students to develop critical thinking skills, for instance, it's not helpful to test them merely on names or dates—on memory skills. In fact most of us have objectives that cover content *and* concept, or knowledge and skill building.

A categorization of learning levels developed by Benjamin Bloom is useful in thinking about what kinds of material and learning we emphasize in our classes. In Bloom's taxonomy learning proceeds from lower order levels of "Knowledge" (simple recall of material) and "Comprehension" (reorganizing information to illustrate conceptual understanding) to middle levels of "Application" and "Analysis," to upper order levels of "Synthesis" and "Evaluation."

If students need to develop skills at analyzing and synthesizing material, then those skills should receive a prominent place, day to day, in the classroom. If your exams only ask for levels of knowledge and comprehension, students will study at those levels and fail to develop higher level skills.

If this seems obvious, it's nonetheless the case that many faculty members test at a lower level than they seem to be teaching. It's also true that sometimes students are surprised by an exam that asks for a kind of synthesis or analysis that was never explained, discussed, or illustrated in the classroom.



From V. L. Clegg and R. E. Owens, *Tips for Writing Tests*, Manhattan, Kansas: Graduate Service Publications, 1983

Construct Tests To Reflect the Classroom Focus

While Bloom's taxonomy is useful in thinking about course construction and the kinds of objectives we need to emphasize, it's less helpful when thinking about the specific questions one asks on exams. R.L. Ebel has adapted Bloom's categories and explained the kinds of questions that relate to his categories. These explanations prove helpful when attempting to identify the focus and content of your exams.

EBEL'S RELEVANCE GUIDE

1. Understanding of terminology (or vocabulary)
2. Understanding of fact or principle (or generalization)
3. Ability to explain or illustrate (understand relationships)
4. Ability to calculate (numerical problems)
5. Ability to predict (what is likely to happen under specified conditions)
6. Ability to recommend appropriate action (in some specific, practical problem situation)
7. Ability to make an evaluation judgment

When we compare Bloom's Taxonomy and Ebel's Relevance Guide, we see that most exams concentrate on knowledge, comprehension, and application. It's important, however, to determine what percentage of each of these you concentrated on in class for the particular topics you've treated. If, for example, you spent 20% of class time on levels of application on two separate concepts, the exam should reflect that fact.

BLOOM/EBEL COMPARED

<i>Bloom's Taxonomy</i>	<i>Ebel's Relevance Guide</i>
A. Knowledge	Terminology Factual Information
B. Comprehension	Explanation
C. Application	Calculation Prediction
D. Analysis	
E. Synthesis	
F. Evaluation	Recommended Action Evaluation

From R.L. Ebel and D.A. Frisbie,
Essentials of Educational Measurement,
Englewood Cliffs, NJ: Prentice Hall, Inc, 1991

Also keep in mind what kind of test is appropriate for the level you want to test at. While true/false and matching questions are appropriate to test for knowledge and even comprehension, they, like multiple choice questions, merely ask students to select information already provided. To test students' ability to produce information, and usually to test for higher levels of learning, you'll need to rely on problem-solving and essay exams. Ideally, you should try to combine levels and types of questions so that students both have to recall and synthesize material.

Prepare Students for What to Expect

Because we focus mainly on content in our courses, we often assume that students clearly understand the best methods of learning this content. But faced with a large volume of information to read and digest, students sometimes flounder, not knowing what to concentrate their energies on. How much should they be memorizing? What are the best ways of synthesizing a variety of concepts? Providing practical advice on study methods before the exam may save you from having to give extensive feedback later.

Of course you can also help students measure their progress before exams by giving other kinds of ungraded feedback earlier in the quarter. Offering review sessions or practice exams can alleviate some students' anxiety while identifying what they need to study. Placing previous exams on reserve also helps clarify expectations. On the most practical level, make sure the students understand the format of the exam—what kinds of questions you'll include and what materials, if any, they need or can use during the exam.

Assessment techniques (which we'll discuss more fully in the winter quarter newsletter) are methods both of seeing how students are doing and also of forcing them to reflect on how well they're learning. For example, asking, at the end of class "What were the most important points covered today?" or "What was the muddiest point in today's lecture?" and having students turn in responses anonymously can help you see whether material is making sense to them.

Feedback after exams is also essential. Students need to understand not only what they did wrong, but why they followed a wrong path. Asking them to do a protocol analysis of their exams, and helping them figure out their errors (either in study techniques or content understanding) will help them develop skills as well as knowledge.

Some Practical Tips

- *Don't wait until the last moment to prepare your exam.*

A simple bit of advice but one that might save you from constructing an unfair or flawed exam: write parts of the test as the course develops and keep a file of potential exam questions. It's more likely that the exam

will reflect the overall course objectives and focus if you regularly pause to consider what's most important.

- *Have a friend, colleague, or your TAs review the test.*

Another set of eyes can help you avoid typos, but also factual mistakes and repetitions. In addition, another opinion about the level and nature of the exam can reconfirm your sense that the test is an accurate and fair assessment of what students should have learned.

- *Ask students to submit possible exam questions.*

By having students prepare questions, you'll get a good sense of what kinds of things students expect on the test. You might adapt or revise some of the questions submitted, or file them away for possible future use.

- *Allow students to comment on test questions.*

Students appreciate being able to explain themselves on questions that they find ambiguous or tricky. You may decide to give partial credit for reasonable responses, or you may notice a difficulty with a question which you hadn't anticipated.

- *Ask for students' feedback about the exam.*

Asking students for anonymous feedback on whether the exam was what they anticipated, and whether it covered the material they expected, or even asking which questions were confusing or difficult, helps you get a better sense of the exam's fairness.

Grades and Grading

There are a variety of ways of grading, from setting up an absolute standard to using a curve. But whatever model you or your department adopts, make sure the students know in advance how they will be evaluated. It is essential to spell out grading policies, in appropriate detail, on the syllabus.

Academic performance—mastery of knowledge and skills—should be the focus of your grade. Other matters, such as classroom behavior, personality traits, attitudes, and effort, shouldn't play a part in the final grade.

It's probably also wise to avoid focusing on grades too often or too strongly. Students are acutely aware of the importance of grades to their futures, and hardly need to be reminded that graduate schools will be scrutinizing their undergraduate records. Encouraging an orientation towards learning rather than towards grades, however, will doubly assist students; they're just as likely to perform at the level of achievement they desire, but in addition they'll probably comprehend and retain information better. They also may hone their ability to continue to learn *how* to learn more efficiently and effectively.

For more information on testing and grading, stop by CTL and pick up a copy of *Teaching at Stanford: An Introductory Handbook for Faculty, Academic Staff/Teaching and Teaching Assistants*. You might also consult our library of teaching resources.

Grading Papers

Grades can only wither away in importance when they cease to be ambiguous and magical. The present system too often allows the student to feel them as judgments based on hidden criteria, judgments which he cannot understand and has little power over. If he is rewarded he feels he did the right things, but if the reward fails he never knows which step in the rain dance he missed.

— Peter Elbow, *Embracing Contraries: Explorations in Learning and Teaching*

Grading papers—as opposed to exams—presents its own set of difficulties. While the feedback we provide students is clearly the most important part of marking essays, we also almost inevitably assign grades. For students to feel that they understand what is expected and how to meet our standards, those expectations should be spelled out. Different disciplines may have different criteria which define the grade we assign, and different kinds of assignments (term papers, lab reports, essay exams) also have their own particular criteria. But in each case it's useful to provide students with a clear sense of what these standards are. Freshmen especially need help in making the transition from the standards used in their high school experience and those being applied in college.

There are clear advantages to developing an explicit set of grading criteria. Not only are students consciously able to work toward a goal, but you as a grader won't agonize quite so much over what grade to assign a piece of writing.

In the English department, Ron Rebholz has developed a set of criteria that freshmen English instructors are given to help them understand how to develop their own grading standards. We offer it here as an example.

The Ron Rebholz statement on grading

An "A" paper must have several qualities. It must have a point: for example, a thesis in an argument, an effect to which a description builds, or a controlling theme in a reflective essay. The point must engage the interest of a reasonably intelligent and informed reader because it has some importance in itself and because the writer gives it a rich development with, for example, cogent logic, illuminating analogies or examples, or abundant and relevant details. The paper must have a clear structure suitable to the point it is making. The syntax must be sound (i.e., no grammatical errors or errors of idiom) and the style lively.

A "B" paper lacks one or two of these qualities, a "C" paper most of them, and a "D" paper almost all of them. I give "N/C" only to papers on which the writer has evidently expended no effort. These qualities are in the back of my mind as I grade, but I do not check for them in a systematic way or deliberately measure the degree to which they are present or absent. Having read the paper twice, I assign it a grade that answers to my impression of it as a whole. In that sense my grading is "holistic."

CTL TA CONSULTANTS

CTL consultants are graduate students with successful TA experience and special training in discussing teaching and learning problems. If you are a TA, you'll meet one of them, or a member of our staff, when you request a videotape session or a quick, in-class evaluation

Derede Arthur	English
Roland Bürgmann	Geology
Hanya Chrispeels	Biological Sciences
Sean Decatur	Chemistry
Sanford Dickert	EE
Michael Gorham	Slavic Languages
Mark Greaves	Philosophy
Rafael Guzman	Petroleum Engineering
Ursula Heise	English
Tracy King	Linguistics
Dennis Kinsey	Communication
David Lowell	EES
Kelly Mays	English
Traci Mann	Psychology
Shelley McConnell	Political Science
Lance Miller	Drama
Linda Price	Industrial Engineering
Marc Sanders	Mathematics
Eric Schocket	English
Jackie Scott	Philosophy
Anne Steinemann	Civil Engineering
Donna Storey	Asian Languages
Jim Tracy	History
Mark Unno	Religious Studies
Michelle Wright	French and Italian
Linda Zimmerman	Art

Consultations for Faculty Members Are Available

Contact Michele Marincovich, Director
or Jack Prostko, Associate Director
at CTL
723-1326

CTL TA LIAISONS

CTL Liaisons are TAs nominated by their departments to help acquaint other TAs, faculty, and lecturers with CTL's services and other teaching resources on campus. Contact your department liaison if you want to know how it feels to be videotaped—or if you have a teaching problem that is related to your department's subject matter or class format.

Aero-Astro	Sally Gressens
Anthropology	Andrea Klimt
Art	Suzanne Wright
Asian Languages	Philip Kafalas
Biological Sciences	Hanya Chrispeels
Chemistry	Susan Shadle
Civil Engineering	Sara Wadia
Classics	Sarah Jones
Communication	Dennis Kinsey
Comparative Literature	Melissa Goldman
Drama	Lance Miller
Economics	Masao Suzuki
English	Tim Wandling
EES	David Lowell
Feminist Studies	Jeannie Alcouloumre
French and Italian	Susan Bree
German Studies	Renée Schell
Geology	Roland Bürgmann
History	Leslie Harris
Industrial Engineering	Michelle Del Tredici
Linguistics	Chris Piñon
Materials Science	Bill Bowen
Mathematics	Marc Sanders
Music	Jean Pang
Petroleum Engineering	Richard Holt
Philosophy	Jackie Scott
Physics	Gideon Friedmann
Political Science	Bari Anholt
Psychology	Traci Mann
Religious Studies	Mark Unno
Slavic Languages	Michael Gorham
Sociology	Liz Boyle