

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Professional and Organizational Development
Network in Higher Education: Archives

Professional and Organizational Development
Network in Higher Education

1991

Course Tests: Integral Features of Instruction

Ohmer Milton
University of Tennessee

Follow this and additional works at: <https://digitalcommons.unl.edu/podarchives>



Part of the [Curriculum and Instruction Commons](#), [Educational Assessment, Evaluation, and Research Commons](#), [Higher Education and Teaching Commons](#), and the [Scholarship of Teaching and Learning Commons](#)

Milton, Ohmer, "Course Tests: Integral Features of Instruction" (1991). *Professional and Organizational Development Network in Higher Education: Archives*. 213.
<https://digitalcommons.unl.edu/podarchives/213>

This Article is brought to you for free and open access by the Professional and Organizational Development Network in Higher Education at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Professional and Organizational Development Network in Higher Education: Archives by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Essays on Teaching Excellence

Toward the Best in the Academy

Volume 2, Number 5, 1990-91

A publication of The Professional & Organizational Development Network in Higher Education (www.podnetwork.org).

Course Tests: Integral Features of Instruction

Ohmer Milton, *Emeritus Professor, University of Tennessee*

The primary function of course tests (and grades) is to rank undergraduates so as to provide admission information to graduate and professional schools. The purpose of this piece is to implore faculty to improve their course tests since those tests are such powerful determinants of learning. Course testing is a very emotional and personal issue and is not much subject to thoughtful considerations.

Please examine the course tests being taken by your college children (or high schoolers) and those given by your colleagues; you will be astounded at the imperfections and limitations therein that inhibit significant learning. One does not have to be an expert in a highly technical discipline in order to recognize many gross deficiencies of course tests. The correct alternative being longer than the incorrect ones is common in m-c questions, and absolutes such as "always" and "never" are used in alternatives and in t-f items. Many students know these and other clues and tip-offs. Negatively worded m-c questions are legion; these tend to confuse the excellent students. Many essay questions are so broad as to demand "shooting the breeze." Students are not helped in learning to think and write clearly by "discuss the need for government with respect to the public good" (one of five in 50 minutes). Another limitation is that too many questions seek only isolated factual information. Do these questions really evaluate course learning?

Furthermore, to what extent is a student's learning distorted by the mad symbol scramble to obtain tenths, hundredths, or thousandths of points (e.g., 3.7, 3.74, 3.749) in order to surpass the arbitrary cut-off points for admission set by many graduate and professional schools? Is it any wonder that nearly 50 percent of undergraduates drop courses for fear of receiving anything other than an A? Some of those who drop will not learn from you or from your discipline. Kenneth Boulding has declared: "Our obsessiveness with arithmetic is the feeling that once a number has been arrived at by a recognized arithmetic ritual something has been accomplished." As the computer increasingly dominates our thinking about the presumed evaluation of undergraduate learning, we are rivaling comedians Bud Abbott and Lou Costello in proving $3 \times 7 = 28$.

Research evidence over a span of at least 50 years documents the notion that course tests are powerful influences over how students study and what they learn; yet very few faculty want to know the research- a strange attitude for scholars. The most compelling evidence for those who will listen and hear comes from students on all campuses. Students warn repeatedly about the effects of course tests upon their learning when they ask these two questions: "Will that be on the final?" and "Will the test be 'objective' or essay?" If the answer to the first question is NO, studying and learning cease for far too many undergraduates. The answer to the second question determines the ways of studying for that test (e.g. memorizing) and the sort of learning that results. Although the influences course tests have upon learning seem to be acknowledged, they are either ignored or denied. A current cartoon depicts a philosophy class beginning to take a test. A student asks, "This 'meaning of life' question-is that essay or true-false?" It is the mad symbol scramble that causes students to ask about the nature of course tests and to modify their studying and learning accordingly. In spite of these apparent relationships between testing and learning, a recent review of research of controllable influences over undergraduate learning (Sherman, 1985) does not even allude to course tests.

At least two explanations for faculty blindness come to mind. First, faculty want to believe otherwise. George Bernard Shaw put it this way: "There is no harder scientific fact in the world than the fact that

belief can be produced in practically unlimited quantity and intensity, without observation or reasoning, and even in defiance of both by the simple desire to believe. . . ." Second, the evaluation of significant learning is tricky, difficult, and inordinately complex, and most faculty have received no enlightenment in "how to." In one of our informal studies, we found that around 75 percent of a faculty sample had never even read about the preparation of course tests, and that nearly 30 percent relied heavily upon intuition when constructing tests.

Faculty committees tinker with grading systems periodically, but little or no attention is devoted to the tests upon which the letter symbols ABCDF are based. Such is an example of attending to peripheral issues rather than fundamental ones. When the tests upon which the symbols are based are faulty, then the symbols must be faulty. Even though a lone faculty member can do little or nothing to alter a grading system, that faculty member can improve his or her course tests.

A few comments about learning will help to put course tests into that broader, more important, and much neglected context. Course tests should be in the service of learning and not in the service of sorting students for society. The term "learning" generally refers to either a process or a product. Process refers to all the different mental activities students must go through as they attempt to learn: general concepts, to reason, to apply, to judge, and so on. Most of the time the term "learning" is used it refers to a product. For example, a high score on a properly prepared test means that learning has occurred. *But learning as a product is an inference.*

There are levels of learning processes ranging from the most simple, or lowest, or easiest (such as recognizing and remembering isolated factual information), to the inordinately complex. We can *infer* the level of learning only by knowing the substance of the test upon which a score or grade is based. And therein is the serious fallacy of relying on a letter symbol in judging the learning of a student. The fallacy is magnified when several symbols from diverse courses are combined into the GPA—that illusion of precision. This statistical ritual of illusion is executed in a variety of ways. But who cares? Some of the isolated factual information sought is trivia because even

trivia ostensibly help in sorting students. Faculty believe widely (but mistakenly) that factual information will be utilized appropriately and more or less automatically at later times. Actually, for most students, this transfer does not occur—a conclusion well supported by research evidence (see Smith, 1989). For personal verification about the limited nature of "transfer of learning" just listen carefully during a general faculty meeting to those characters from disciplines other than your own.

Institutional recognition of the need to assist faculty in their preparation of course tests will continue to be postponed by the burgeoning "assessment" rage—endeavors remote from the center of everyday teaching/learning. Already bureaucracies are being formed and at least one university system has created the position of vice president for assessment. It will be a long time, if ever, before any of the questionable results of the assessment movement will reach faculty members and assist them in influencing learning. Here is another instance, alas, of dealing only with the surface issues.

Suggestions: A cartoon of the early 1980's has one professor saying to another: "Hoo-boy! If we're really looking for better answers. . . maybe we should start asking better questions." Here are some suggestions individual faculty can take toward that end:

1) Exercise as much care in writing each test question as you do in other sorts of writing—statements of academic policy, for example. Many students read test questions more carefully than they do any other material; thus they spot the flaws and confusion results.

2) Cease using the term "objective." M-C questions are *not* objective. Those questions do not come from thin air—especially those in manuals accompanying textbooks. A person decides to question this rather than that and *then* writes - subjective processes. The term "objective" misleads both students and the public. Correct students when they use "objective."

3) When using m-c questions, design them to tap higher-order thinking processes. (For illustrative m-c items, see, for example, *Constructing Achievement Tests* by Norman E. Gronlund, Prentice

Hall, 3rd edition, 1982, chap. 4.)

4) Limit the scope of essay test questions, or else they become exercises in "shooting the breeze" that do not help in promoting clear thinking. (See Chapter 5 in Gronlund). Another way to limit the scope of a question is to use more specific words like "compare," "contrast," "criticize," and "explain." Avoid "discuss;" that word is used quite ambiguously.

5) Ask a colleague to review all questions prior to their use: "Are the questions clearly written?" "What level of learning does each question tap?" "Do the m-c question alternatives contain tip-offs?" "Are there unnecessary negatives in the m-c questions?"

6) Join forces with other faculty and push the administration to provide assistance with this time-consuming but powerful teaching/learning tool-testing. As a beginning, review the "Board of Examinations" program used at the University of Chicago during the 1930s and 1940s (Bloom, 1954).

References

Bloom, B.S. (1954) Changing Conceptions of Examining at the University of Chicago. In *Evaluation in General Education*, P.L. Dressel, Ed., Dubuque: IA: Wm. C. Brown, Co.

Milton, O and Eison, J. (1989). "Better Course Examination Questions: Guidelines." Knoxville: Learning Research Ctr., U. of Tennessee, duplicated.

Milton, O., Pollio, H. & Eison, J. (1986) *Making Sense of College Grades*. San Francisco: Jossey-Bass.

Sherman, T.M. (1985). "Learning Improvement Programs: A Review of Controllable Influences." *J. of Higher Education*, 56 (1), 85-100.

Smith, M. (1989). "Why is Pythagoras Following Me?" *Phi Delta Kappan*, February.