



Faculty Forum

From the Faculty Center for Teaching Excellence

WESTERN CAROLINA UNIVERSITY

Vol. 2, No. 2

CULLOWHEE, NORTH CAROLINA

September 1, 1989

Designing an Effective System of Evaluation

One of the most difficult tasks for any teacher is designing an effective system of evaluation. An evaluation of a student's performance is not simply an act at the end of a term or grading period. It is something which should play an integral part in the learning process. I suspect that for most of us, the method we use to evaluate our students is our greatest shortcoming as educators. Here are some questions I think we need to consider whenever we are designing a system for evaluating student performance.

(1) Does our system emphasize the completion of the course objectives by the end of the course or does it emphasize completing course objectives before some arbitrary date, such as the first or second exam? Imagine that a student eventually completes an objective on the final exam but failed to complete the objective on an earlier exam. Is the earlier score valid? Has the student failed or simply failed to complete the objective according to an implied timetable? As I ask my students, if a cake recipe calls for 4 eggs and you use one rotten egg, can you throw away 1/4 of the cake when you are finished? If you average grades to arrive at a final course grade, does this mean that one of the scores in your evaluation is invalid? If so, then probably so too is your final evaluation. Students learn at different rates and these differences are probably legitimate. If students are working hard and just need more time to learn, shouldn't we give them the extra time? It's difficult, of course, to know when students are working hard enough to fulfill their end of the bargain, but it's not impossible to ascertain their level of effort. This makes our evaluation process more complex and difficult, but it also makes it more fair and probably more effective as a stimulation of the students' learning.

(2) What role does luck play in our evaluation systems? Each semester I administer a "blind" 10 question T/F exam to my education majors, asking only for 10 T/F answers, without providing the questions. Invariably, some student will get 8, 9, or even 10 answers correct. A multiple choice question in which there is only one good distractor is no better than a T/F exam. Would we all agree that luck should not play a significant role in our evaluation systems?

STEERING COMMITTEE: FACULTY CENTER FOR TEACHING EXCELLENCE

Scott Higgins, Chair	(7113)	Judith Clauss	(7295)	Frank Prochaska	(7260)
Maynard Adams	(7272)	Michael Dougherty	(7207)	John Wade	(7401)
Jim Addison	(7264)	Wilburn Hayden	(7112)	Jim Wallace	(7244)
J.C. Alexander	(7436)	Mary Anne Nixon	(7401)		

CENTER STAFF

Ben Ward, Director	(7196)
Jane Hall, Fellow	
Terry Nienhuis, Fellow	
Charles Stevens, Fellow	

(3) If we analyzed our last exam would we find that the bulk of our questions address memorization skills instead of higher forms of human thinking? The students can memorize fairly well when they want to or they probably would not be in college. To be successful in our disciplines in the real world, does one need more than memorization skills? To what extent and in what ways does one need to be able to use higher forms of thought and reasoning in the world of work after formal schooling? Shouldn't our evaluation systems reward higher order thinking much more than memorization? Here we run into a problem of time. An evaluation system that measures simple mastery of a finite set of information bits is easy to construct because we only need to score and add up numbers. An evaluation system that measures the student's ability to think is much more complicated and requires much more time for construction and grading [see Hanna and Cashin (1987), "Matching Instructional Objectives, Subject Matter, Tests, and Score Interpretations," available from Ben Ward at the Center].

(4) When we determine a student's final grade do we let the calculator do the thinking, or is the calculated average merely the starting point in our analysis? We know that the "mean score" is but one measure of central tendency and not always the best measure, but how often do we put this belief into practice when we create a letter grade for a student's final evaluation? Students learn best when they feel that our grading is fair and rational rather than arbitrary and capriciously subjective. Isn't it our job to convince them that they can trust our grading?

I believe that if we examine the situation, almost everything we do to make our evaluations easier for us, such as administering short answer exams and using the mean score exclusively, is performed to the detriment of a valid evaluation. Determining what a person understands should be one of the most difficult and complicated tasks that we face. Steps we take to make it easier and less complicated may only serve to invalidate our efforts to educate effectively.

Charles Mitchell, Mathematics

Editor's note: Charles Mitchell taught at WCU from 1986 through 1989. He is now teaching at Western Illinois University