
Testing : A Critical Part of Good Teaching

Khedair S. Al-Khedair

King Fahd University of Petroleum & Minerals
Dhahran

Teachers and academics are always involved in testing. This function takes a great portion of their time either for making a test or testing students for measuring their achievement. Tests always involve a great deal of preparing, administering, correcting and assigning grades. The notion of testing has always been a fact of life to academia since testing measures and evaluates students' learning. Tests are constructed in different formats to meet different functions. There are some general ability tests such as the Stanford Binet Test and Achievement Tests which are generally given after a period of teaching and administered by teachers to measure how much students have learned from the classroom instruction. There are also syllabus-based and non-syllabus-based performance tests. The distinction between the two is useful, because it enables us to capture the essence of the difference between achievement tests and general performance tests (Dunseath, 1981). The purpose of the achievement test is to find out how much of a syllabus or part of a syllabus a student has learned and understood. In this sense, the achievement test is an extension of the mastery test, which is used to test control of single items or patterns.

These tests are set at different levels of difficulty, depending upon the ability of the students in the class or how well the teacher thought certain topics were covered or taught. Tests are also written in different formats. There are objective multiple choice tests, essay tests, fill-in-the blank tests, completion tests, matching tests, true/false tests and free response

tests. Each test has its advantages and drawbacks depending on the tester's viewpoint with regard to its practicality in relation to time, administration and cost.

Classroom test writing has undergone significant development during the last ten years as a result of deep interest in writings on behavioral objectives, where (1) more emphasis is put on the idea that students have to be tested on a preset objective; and (2) testing has to cover the information taught in class.

Testing as an assessment technique serves many purposes. According to Emans (1990), testing is considered to be an integral part of teaching and learning experience of the students, the questioning and answering process in class and through the observation of students. Another function of assessment is providing a feedback to the instructor and the students about the latter's progress early enough, so that needed changes and the instructional efforts can be made. On the other hand, summative assessment determines the degree of achievement of major outcomes of a student's course of study.

Testing provides a feeling of professional assurance for the teacher as well as an indicator of psychological relaxation to the students about their learning capabilities. Testing could be considered in terms of what we learn about the needs students usually bring to the classroom. The need to meet challenges and overcome difficulties is a basic and long standing psychological motive. Tests, like athletics or artistic

competition, meet this need for many students. A good teacher is expected to create a test that strikes a balance between meeting students' need to feel secure and meeting the need to strive for challenges. The absence of challenge for students in any given test is as unsatisfying to them as making an unnecessarily harsh or unfair one. Testing also gives a teacher some awareness of class norms and of each individual's standing or progress with respect to these norms.

Regarding the relationship between teaching and testing, Hughes (1990:2) says "that there is a partner relationship between teaching and testing." This latter opinion is totally supported by Carline (1991:22), who states that "Good Evaluation is a Critical Part of Good Teaching."

When writing tests, the teacher has to be able to deal with these essential questions:

- (1) why am I testing?
- (2) what am I testing?
- (3) what results am I getting?

The first question comes to the teacher's mind because of the complaints students voice among themselves and sometimes with the teacher, which are more pointed, but they fall within this category. Some of these complaints deal with the irritation of small and trivial quizzes, confusion over just what the teacher wants, objectives for what the teacher accepts as perfect or poor answers and a headache or sick stomach both before and after going 'down the tube'. It

Types of Test

Type of Behavior from Objective	Essay	Fill-in-the blank	Completion	Multiple Choice	Matching	True/False	Free Response (building, design, etc.: not paper and pencil)
State	x	x	x				
Identify	x	x	x	x			
Discuss	x		x				
Define	x	x	x				
Select				x	x	x	
Discriminate				x	x	x	
Solve	x	x	x	x		x	x
Develop	x		x				x
Locate	x	x	x	x	x	x	x
Construct	x	x	x				x
Generate	x		x				x

Fig. 1 (adapted from Walter Dick-the systematic design of instruction-1978:83)

is because of these responses that we have to explore the first question above: **why am I testing?**

Eble (1967) in his book *The craft of teaching* talks against the practice of using tests mainly to assign grades. He says "that it is indefensible practice, except where grades are finally established in the student's minds as a measure of learning...even then it is indefensible unless the grade and the test point out what has been learned, what remains to be learned, and

what is vaguely comprehended or wholly understood. Keeping in mind this initiated sense, then all tests are diagnostic, telling the students some very specific things about where they stand in regard to the development of skills or the acquisition of information. If we argue that this diagnostic function is legitimate then it would be of a great assistance to the teacher in preparing more useful, valid, and focussed tests to build specific abilities and to adjust them to

individual needs" (Eble, 1976 : 102).

The second question that the teacher has to deal with in constructing a test is: **what am I testing?** What are the objectives I would like to test, high objective, medium objective or low objective? It is the general practice that teachers often choose questions solely in terms of class size, using multiple choice test for a large class, short-questions for medium size classes and essays for small classes (McKeachie, 1968). It seems that class size is an important factor in deciding the type of test the teacher will give, but it is more important that the teacher should take into consideration the educational goals and course objectives when writing their tests. In spite of the fact that class size may dictate the type of test the teacher will write, it is advisable to look at figure I as a guideline for teachers when they are thinking about writing their tests.

The third important question in testing is: **what results am I getting?** The answer to this question lies in what the test reveals and what it does not. What the test result should not tell the teacher is what they most often do tell them - that there are bright students and below average ones. It seems to be hard to avoid such off-hand judgements, but there are certainly ways of rendering them relatively painless or harmless. One way is not only sorting the test items into what's hard to answer from those which are easy for everyone, but to look at tests from a diagnostic point of view and try to discover the best points of each student's performance. Feedback from the whole class can focus attention on known and unknown material, sometimes allowing opportunities to work on individual student learning in specific ways. It is often concluded that test feedback is as important and as worthy of care, intelligence and imagination as making up the test itself.

In developing and writing tests, dealing with certain concepts such as test validity and test reliability is unavoidable. Test validity means that the content and the procedures of a given test are considered to be a good basis for decision-making and the test is considered a good instrument for measuring what is supposed to be measured. Test reliability on the other hand means that we have to construct, administer and score tests in such a way that the scores actually obtained on a test on a particular occasion are likely to be very similar to those which would have been obtained if it had been administered to the same students with the same ability, but at a different time. The more similar the scores

would have been, the more reliable the test is said to be (Hughes, 1990). In order to increase the reliability of any test a teacher should keep the following in mind:

1. increase the length of test with questions of the same standard;
2. reconsider rewriting or replacing the bad questions;
3. be sure about the circumstances of administering the test;
4. instructions of the test should be as clear as possible to the testees.

Kuder Richardson's formula 20 has been developed to provide a statistical calculation of reliability as follows:

$$K = \frac{n(V - \sum x y)}{(n - 1)V}$$

Where K = the reliability estimate
 n = number of items in the test
 V = the variance of the test results
 x = the proportion of correct responses to the item
 y = the proportion of incorrect responses
 and $\sum x y$ = the sum of the products of x and y for each item. (Baker 1989 : 61)

As the teacher writes a test for the purpose of measuring the students' level of achievement, then dealing with grading and grade assignment becomes unavoidable.

The grade is a judgemental statement that a teacher passes on a completed test taken by a student. This judgement is generally instituted by a letter (A, B, C, D and F) or a figure (0--100) depending on the system adopted by the institution to which the teacher belongs. This letter grade or figure/score is usually qualified by a qualitative adjective such as excellent, very good, good, average, below average or poor. No matter what grade a student deserves or gains, it should be based on certain instructional goals that take into account the content of instructions and the cognitive complexity of outcome and data collection about the student from the beginning of the course.

Assigning a grade is a process that all teachers find distressful and painful. To some teachers the joy of teaching is minimized and faded when they think about

assigning grades. It is not surprising that some authorities in the field of measurement such as Terwillzer (1988:15) argue that "Assigning grades to students is undoubtedly one of the most distasteful aspects of teaching. If pushed, most teachers will state that the assignment of grades is, at best, a necessary evil that has little to do with the task of teaching."

As a teacher, you may find yourself obligated to assign a grade. Administrators are always in need of evaluation and feedback on the educational process and students' learning achievement. In case you are in such a situation where there is heavy emphasis from the administration on students' grades, it is suggested that the following points on assigning final grades may be helpful and useful:

1. Following a strategy of continuous assessment of your students, do not depend on only one final exam. It is recommended that each test should have a grade and students should be aware of their grades on that test. The average of these tests should contribute to the final grade of the course.

2. Design a system of grading that recognizes and rewards the effort of the students all along and not just at the end.

3. Be an encouraging agent for all students, even those whom you feel are poor or below average.

4. When calculating final grades, perform each computation more than once and keep the marginal cases together. Teachers have to apply their wisdom in assigning grades for border-line students. Wisdom and fairness are the most essential factors in grading.

5. It is suggested that you should avoid the state of mind that pre-sets your thinking about quota such as no more than 15% to get As or 30% to get Bs.

6. If the institution requires the teacher to post the students' grade, do it in such a way that the identity of each student is protected.

7. The teacher should avoid change in the assigned final grade unless such a change is warranted. If you are sure about your calculations and fair in your judgement, don't respond to a student's requests for higher grades.

Teacher	Course No.	Course Title	Grade Distribution
Dr. Smith	101	ENGLISH	10% First Test 30% Mid-Term 10% Second Test 50% Final Test

TABLE 1

TEST	MAXIMUM GRADE	DESIRED WEIGHT
Test 1	50	10%
Mid-Term	100	30%
Test 2	50	10%
Final Test	100	50%

Since we are advocating continuous assessment for students' work during the course, you may find yourself assigning percentages for each test that could measure the students' progress through the course. Let us say that you have decided to give two tests additional to the mid-term and the final, and you have decided that the first test will count for 10%, the mid-term will count

for 30%, the third test will count for 10%, and the final test will count for 50%. In order to arrive at one single final grade at the end of the course, you have to find a way of weighing these exams that will help you to arrive at that "grade" that a given student should get for the course. The following illustration and formula may be useful in this regard.

TABLE 2

Calculation of the final grade based on desired percentage weight.

Course Title: ENGLISH

Course # 101

Section # 4

Teacher: Dr Smith

Students	Test # 1-10%		Mid-term Test -30%		Test # 3-10%		Final Test 50%		Total Score	Final Grade
	Max.	Achieved	Max.	Achieved	Max.	Achieved	Max.	Achieved		
	Score	Score	Score	Score	Score	Score	Score	Score		
No.1	50	30	100	65	50	45	100	80	75	C+
No.2	50	40	100	70	50	35	100	70	71	C
No.3	50	35	100	75	50	25	100	65	67	C-
No.4	50	25	100	55	50	40	100	85	72	C
No.5	50	45	100	40	50	20	100	75	62	D
No.6	50	45	100	80	50	45	100	85	90	A
No.7	50	40	100	65	50	35	100	75	72	C
No.8	50	25	100	35	50	20	100	50	44.5	F
No.9	50	35	100	85	50	45	100	90	86.5	B
No.10	50	30	100	55	50	35	100	70	64.5	D

In order to calculate the weighting for these tests, we have to multiply the student's achieved score by the percentage given to each test and divide by the maxi-

mum marks allocated to each test. Table 2 will provide a total picture of the calculation of the weighted score for each test and the final grade assigned to each student.

The formula for such a calculation is as follows:

$$\frac{\text{The achieved score} \times \text{the desired percentage weight}}{\text{The maximum score assigned for the test}} = \text{Weighted test score}$$

In conclusion, teachers should also keep in mind the list of "never." That list has been adopted and expanded from that given by Lowman (1984).

1. When you agree with your student about a given type of a test, you should stick to this test. Do not change it, because you may create more anxiety to students in addition to what they are experiencing about testing in general.

2. When you finish your test instructions to students, don't give conflicting instructions. This will contribute to the wasting of your time as well as the students' time.

3. Whenever using a multiple choice test, never put two correct answers for any question.

4. If you are correcting an essay question, you should read all the pages written by the student. Never

skip any page of the student answer.

5. Avoid giving similar grades for every student on an essay question.

6. Avoid labelling students from the first exam you gave and deciding that this student is an A student and that one is a C student.

7. Never fail to give an indication during a course of the grade that a student is likely to achieve.

8. Never tell a given student that you had given an 'A' grade but the department chairman objected to it on the grounds that too many A's had been given in a particular course, and, as a result, the grade was lowered from A to B.

9. Avoid giving a student a lower grade for suspected cheating without telling the student himself and the department chairman.

References

- Carline, J.D. (1991). *A Good Evaluation - A Critical Part of Good teaching*. A short course offered at K.F.U.P.M., November 23-27.
- Dick, W. & Carey, L. (1978). *The Systematic Design of Instruction*. Illinois : Scott, Foresman & Co.
- Dunseath, K.J.C. (1981). *A Typology of English Language Tests*. M.A. Thesis, St. John's College, Cambridge.
- Eble, E. (1967). *The Craft of Teaching*. San Francisco : Jossey & Bass.
- Emans, L. (1990). *Understanding Undergraduate Education*. Vermillion, South Dakota: University of South Dakota Press.
- Hughes, A. (1990). *Testing for Language Teachers*. New York: Cambridge University Press.
- Lowman, J. (1984). *Mastering the Techniques of Teaching*. San Fransisco : Jossey & Bass.
- McKeachie, W.J. (1986) *Teaching Tips. A Guidebook for the Beginning College Teacher*: Lexington, Toronto : D.C. Heath. & Co.
- Terwillizer, J.S. (1988). *Classroom Assessment Research and Training* (Paper). University of Minnesota.