

Teacher-Made Exams: Part I

Gary A. Holt
Kay E. Holt

A PHILOSOPHICAL ORIENTATION TO TEST WRITING

Teacher-made tests are the most frequently used device for collecting information about students and are the common way by which most educators evaluate their students. In fact, testing is such a common practice that educators may not give it the consideration it is due. They often reason that because exams are taken by classes of students all is probably fair. After all, if one student can perform well on a test, others should be able to do so. Thus it is assumed that students who do not perform well are simply not applying themselves, and sometimes this is true. Yet it is also possible that students perform poorly on exams because the test instrument has been poorly designed. Furthermore, even properly designed tests can sometimes fail to evaluate students accurately.

The test itself should be used as a learning tool. This requires that tests be designed appropriately to achieve desired results and to ensure validity and reliability of the measurement process. There are two inescapable realities regarding tests that should be in the mind of every educator as (s)he prepares tests.

Gary A. Holt, M.Ed., Ph.D., is Assistant Professor of Pharmacy at the School of Pharmacy, University of Wyoming, P.O. Box 3375, University Station, Laramie, WY 82071-3375. Kay E. Holt is a doctoral student in the School of Education at the University of Wyoming.

This article is the first in a series on teacher-designed examinations contributed by Gary and Kay Holt.

First, there is an art to writing tests, just as there is an art to teaching in general. Learning how to write good tests is not a simple matter. It is not a skill that can be learned solely from textbooks. Rather, it takes time and practice to develop these skills. Even so, there are many guidelines that can facilitate the effort.

A second reality of test writing is that there is no perfect test. Despite the efforts one devotes to test writing, most questions can be interpreted in numerous ways, particularly by intuitive students. Thus failure to answer a question correctly may not mean that the student does not know the answer; it may simply mean that the student did not understand the question. Failure to understand questions is not necessarily equivalent to lack of preparation on the part of the student because even well-prepared students can sometimes misinterpret test questions. In fact, highly prepared, intuitive students are quite likely to misinterpret or to become confused by certain types of questions (e.g., true/false, certain multiple choice, etc.).

A concern similar to this latter point involves the use of testing techniques that penalize students for guessing. These include testing formats, such as right-minus-wrong questions, where students lose more points by missing a question than they can gain by indicating a correct answer. The assumption underlying this testing approach is that the questions are so clearly written that no misinterpretations can occur. Accordingly, the student either knows the correct answer, or (s)he does not. The student is instructed not to answer a question if (s)he is not certain about the answer.

The problem with this latter approach is that virtually any test question is subject to some interpretation. Again, this interpretation need not mean that the student is guessing or that (s)he is unfamiliar with the material. Testing formats that penalize students for guessing are probably not rational, since guessing on exams really cannot be eliminated. Furthermore, a reasonable argument can be made for potential benefits that may be realized by students from learning to make educated guesses. It is a process by which students deduce new information and ideas from information they already know. Conceivably, educated guessing could help students to develop heuristic thinking skills that could be of tremendous benefit to them in their professional careers. Indeed, educated guessing is a requi-

site component of scientific research, so it is ironic that heuristic thought processes are actually discouraged in academic environments. Because it is impossible to design tests that discriminate between uninformed guesses and heuristic guessing, it is far more beneficial, from a learning perspective, to allow for heuristic guessing on exams.

Problems with testing accuracy also exist because some students simply perform better on certain types of tests than on others. For example, some students perform much better on objective tests (e.g., multiple choice, true/false, matching, etc.) than on essay tests. In order to ensure measurement accuracy, it is important for educators to recognize the differences among the various types of test items (to be discussed in a later article in this series). Teachers may want to vary the types of test items throughout the school year so that students are not handicapped by specific testing techniques. If one test design is being used throughout the semester, the teacher may want to allow students opportunities to write papers or to participate in other class projects to offset the effects of certain test designs. If this situation is not taken into account, educators may often be measuring students' ability to take certain types of tests rather than their knowledge and understanding of course content. For example, most objective tests are based upon memorization of course content. Memorization is considered representative of the lowest cognitive skill (1-6). Yet contemporary education tends to emphasize this skill to the exclusion of higher-level cognitive skills, such as comprehension, analysis, synthesis, and evaluation. Students who have an affinity for memorization may perform well on objective examinations even though their actual understanding is equal to, or even less than, that of some students who perform poorly.

Up to this point, we have emphasized that educators cannot assume that any test is an accurate reflection of student performance. Educators are prone to blame students for academic failures even when the test or the educator may actually be the primary problem. Unfortunately, in our society there are many poorly written examinations, and many college educators are not motivated to improve either their test-writing skills or their other classroom competencies.

Insight into appropriate test-writing skills must necessarily begin with a consideration of the function of grades and testing and even a consideration of what tests are supposed to be. In particular, testing allows for the evaluation of the student and of classroom instruction relative to course and program goals and objectives. Tests are widely viewed as a means by which student mastery of course content can be determined and as a measure of academic progress. But they can also help to identify gaps in student learning and can serve as a tool for indicating the degree to which classroom objectives have been accomplished by the teacher (7). Thus it is important to realize that exams reflect not only student performances but also those of instructors. Because test scores and grades provide feedback to the academic institution regarding its effectiveness, exams can be used as a basis for evaluation and revision of both courses and programs. In order to accomplish objectives, grades must measure *actual* student performances. This means that tests and grades should be accurate. Nothing should be allowed to detract from measurement accuracy.

Exams should not be used to penalize or punish students. This is, unfortunately, a common practice, particularly where pop quizzes are concerned. Such a practice generates negative classroom attitudes, contributes to poor teacher/student relationships, and calls into question whether or not such a practice is an accurate measure of a student's academic performance.

Grades are also used to classify students, to serve as a basis for student guidance and counseling, and to motivate students. Motivation, in particular, is a somewhat controversial function for grades because it can detract from learning for its inherent value. Our educational system has conditioned most students to be overly grade conscious, often at the expense of regard for the value of learning, knowledge, and understanding. Accordingly, students may opt for easier courses and easier instructors in an effort to maintain desirable grade point averages, even though harder courses and instructors may provide greater learning opportunities. Students also may fail to give due consideration to important concepts if they perceive that only one or two questions on a test will refer to these concepts.

Stated simply, tests are nothing more than a group of questions or tasks to which a student is to respond (2). From this process, sys-

tematic information is gained that is presumed to be representative of educational or psychosocial traits or attributes (6). Yet tests must be seen as more than a collection of questions or a traditional exercise with which students must contend. Those tests not designed with the full insight and expertise of an informed educator probably do not serve a useful purpose and may distort both the student's actual academic performance and the performance of the educator and the institution.

There are, conceivably, almost an infinite number and variety of questions and test designs that could be used to assess student progress or measure desired student traits. A test may be viewed as a representative sampling of all possible questions and tasks of appropriate difficulty that are related to the trait(s) being measured by the test. But a test must be a sample by design and not a sample by convenience. Each question should appropriately address relevant course content. Questions should be avoided that simply require students to recall trivial or meaningless material on exams. Instructors may do this in an effort to create more difficult exams. This is inappropriate, and it threatens the validity of both educational and testing efforts. For example, many pharmacy educators question the value of having students memorize hundreds of chemical names and structures solely for the purpose of recalling them on a test. Because the memorization of such information is not particularly useful to pharmacists in practice, one may well suggest that tests should address more relevant concepts that are more meaningful and useful. Furthermore, tests should not be used merely as memory exercises.

Finally, testing is a form of communication between the student and teacher. The testing effort presumes an ethical and responsible obligation on the part of the teacher and a willingness to cooperate on the part of the student. As with all forms of social interaction, a relationship of mutual trust, respect, and understanding should be developed if the testing effort is to achieve desired goals and objectives. Because the student is expected to devote considerable time and effort to test preparation, (s)he has the right to expect benefit and value from the testing effort (6). Failure to establish a relationship of trust and respect will detract from testing efforts and seriously threaten the validity and reliability of the test.

The three basic factors requisite for the development of appropri-

ate tests are an ability to use and manipulate the English language, a knowledge of the subject matter, and an understanding of tests and their uses (7). It is assumed that most educators possess the first two traits. Yet most pharmacy educators have had minimal coursework, if any, in educational concepts. Thus it is appropriate to consider and understand the purposes and techniques for testing and grading. Such a philosophical inquiry goes far in promoting rational and accurate evaluation techniques that benefit all involved.

TEST ADMINISTRATION

It is not sufficient merely to write an appropriate examination; the test must also be administered appropriately. Table 1 summarizes points of test administration that should be considered.

TABLE 1. Test Administration Guidelines (1)

1. Make certain that students understand test directions and classroom policies during the exam.
2. Develop a policy for handling questions and problems during the exam. Convey this policy to the students along with other test instructions.
3. Remind students of policies and consequences regarding academic dishonesty.
4. Make certain that the physical environment of the classroom is conducive to test-taking efforts.

First, educators should make every effort to ensure that students understand the test directions. Unfortunately, some educators write confusing directions or actually hide some directions within the test to see if the students will notice them. This is not appropriate because, once again, such tests are more accurately a reflection of the students' ability to take tests than of their actual knowledge and understanding of course content. In order to ensure that the exam instructions are understood, instructors should read the directions with the students. If there is still concern about certain questions, the teacher can observe several students to see if the exam is being completed correctly.

Teachers should establish a format for handling problems that

arise during the test period. This includes a procedure for answering student questions during exams. Many educators do not allow students to ask questions during the exam period. The rationale is that students who ask questions may inadvertently be given additional information, giving them an advantage over others. It can also be argued that such a policy is too restrictive. Students often spot problems or discrepancies on exams that need to be brought to the attention of the instructor. Additionally, many bright, intuitive students may actually be penalized by such a policy, since there is not opportunity for clarification of questions that may be confusing or ambiguous.

Any policies selected should be conveyed to students along with other test instructions. If questions from students are allowed, it may be least disturbing for students to raise their hands so that the instructor can come to them. This would also decrease the opportunity for the student to observe the papers of others while coming to the front of the room. Other instructors prefer that students bring their papers to the instructor so that they are more removed from the rest of the class for discussions and questions. In either case, the adoption of a policy minimizes confusion and distractions during the exam so that all students are better able to concentrate.

Educators are also well advised to adopt classroom policies regarding cheating. Even though most colleges and universities have regulations and policies regarding academic dishonesty, at least minimal precautions should be taken by individual instructors to reduce the opportunity and temptation for dishonesty. Table 2 summarizes some common precautionary measures.

TABLE 2. Teacher Measures to Reduce Academic Dishonesty During Examinations

1. Observe students frequently during the exam.
2. Move throughout the room periodically during the exam. (NOTE: At the same time, instructors must make every effort not to distract students.)
3. Remain in the room during the exam.
4. Provide alternate forms of the exam with the questions rearranged so that students sitting next to each other do not have identical forms of the test.

The adoption of formal policies serves as a reminder to the students that academic dishonesty will not be tolerated, either by the institution or by the instructor, and that the teacher will be attentive during the exam. Teachers who allow dishonesty penalize honest students, actually promote cheating, encourage undesirable behaviors, waste the time of both students and classes, and endanger the validity of the testing effort. Policies should be adopted to ensure that all students are being tested under the same circumstances (1).

A final consideration of test administration involves the physical classroom environment. In particular, any situations that detract from the students' ability to take an exam should be eliminated if at all possible (1). This can include a consideration of noise, lighting, seating arrangements, or other physical parameters. Physical comfort can also include individual instructor policies regarding smoking or beverages during the test period, if the institution does not already have such a policy.

By establishing policies for test administration, instructors will likely find that only minimal problems and distractions disrupt their examinations. Most educators find that such policies work well for them, and students find the examination period more secure and successful because test-taking efforts are enhanced.

REFERENCES

1. Mitchell RJ, Westerman JE. *Evaluation in the classroom*. Dubuque, IA: Kendall/Hunt Publishing Co., 1977.
2. Ahmann JS, Glock MD. *Measuring and evaluating educational achievement*. Boston: Allyn and Bacon, Inc., 1971.
3. DeCecco JP. *The psychology of learning and instruction: educational psychology*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1968.
4. Good TL, Brophy JE. *Educational psychology—a realistic approach*. 2nd ed. New York: Holt, Rinehart and Winston, 1980.
5. Lindgren HC. *Educational psychology in the classroom*. 6th ed. New York: Oxford University Press, 1980.
6. Sax G. *Principles of educational measurement and evaluation*. Belmont, CA: Wadsworth Publishing Co., Inc., 1974.
7. Bergman J. *Understanding educational measurement and evaluation*. Boston: Houghton Mifflin Co., 1981.