Teacher-Made Exams: Part 2

Gary A. Holt
Kay E. Holt

INTRODUCTION

The ability to write tests that are valid and accurate is difficult for many educators, and it is not a skill that is easily taught. There is no simple way of learning how to develop good tests. Primarily, it is a skill that most educators acquire with experience (1). Even so, there are guidelines for test construction that can facilitate test writing efforts and ensure that the test will achieve the instructor's objectives.

TEST PLANNING AND WRITING

As with most educational activities and efforts, test outcomes are usually best when appropriate planning has taken place. As indicated by Table 1, planning involves more than merely setting a date for the exam and having a secretary type a list of indiscriminately selected questions.

It is important to keep the purpose of the test in mind. Mastery tests are usually used to ensure that students have mastered desired materials and skills (1, 2). Because pharmacy programs involve the mastery of materials that are requisite for later professional practice, thought should be given to an appropriate level of difficulty. This level of difficulty should reflect program goals and course ob-

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.

TABLE 1. Points to Consider When Planning and Writing Examinations

1. Keep the purpose of the test in mind. Design a level of difficulty that is appropriate for the level of students in the course.

2. Questions should reflect course objectives. Learning objectives for the course should be provided to the students so that they know what learning outcomes are expected of them. Tests should reflect these objectives.

3. Decide upon the type(s) of questions that will be used on the test.

4. Plan for an appropriate test length.

5. Tests should stress content validity.

6. Develop a table of specifications to help balance the test.

7. Design test questions that are clear, concise, and easily understood. Avoid trick questions.

8. Avoid copying test questions directly from material in the text or other course materials.

9. Prepare tests several days in advance. Review questions prior to giving the test to check for clarity.

10. Write each test question on an index card in order to create a test question pool. This facilitates test preparation efforts and allows for convenient review of questions periodically.

11. Tests should reflect the course, rather than the course being geared to tests.

jectives relative to the materials that have been taught (3). Tests that are too difficult not only affect student performance but also threaten test validity. Overly difficult exams can lower student morale and detract from student efforts. In such a case, the test can become a measure of the students' ability to endure a particular type of test rather than a true reflection of knowledge.

Some educators incorrectly rationalize that when students are enrolled in a professional program, any level of difficulty is appropriate. This simply is not true. For example, question language, question types, and concepts that are appropriate for doctoral students usually are not appropriate for undergraduates or even for Doctor of Pharmacy or master's students. Difficulty levels should reflect the objectives of the course and the particular program involved.

When test questions reflect course objectives, the test has greater validity. Additionally, lists of the course objectives should be provided to the students as a study guide for the course. Many educa-

tors object to this approach (competency-based education) because they think that the objectives list essentially provides the students with a preview of forthcoming exams. Instructor concern is one of grade escalation. But this concern, which is largely based on the bell curve concept, is an unreasonable concern when education is the true objective. It is also likely that many educators avoid using objectives because they perceive it to be too much work or because doing so denies them the opportunity to conveniently manipulate grades. Theoretically, if educators teach well and students learn what they have been directed to learn, there is really no reason why the overwhelming majority of students should not earn all A's and B's. In fact, it would seem inherently desirable that all students in the health professions should perform well academically if they are to serve the public well. Course objectives should relate to the students what the educator considers to be important information. Students have a right to know what the objectives of the course are and what learning outcomes are expected by the instructor. The educator has a responsibility to define these objectives, if only to give appropriate thought to information priorities. The notion that any and all information is fair game is an inappropriate educational objective. The test should reflect course objectives to ensure that the students have actually learned the information requisite for functioning properly as health care practitioners.

It is also important to decide upon the types of test questions that will be used (e.g., true/false, multiple-choice) (2). All too often, insufficient consideration is given to the types of test questions selected. The advantages, disadvantages, and methodologies for preparation of various question types will be discussed in subsequent parts of this series. At present, suffice it to say that question types can vary significantly in ease of preparation, validity, and suitability for soliciting specific types of information. Additionally, students can respond quite differently to different types of questions that cover the same material. This is of particular concern because the objective of a test is to accurately measure student mastery of course content and not to measure the students' ability to take certain types of tests.

To offset problems that can arise from using one type of question, some educators prefer to use a variety of test items with a

balance among types. Thus, if students are weak with one type of test question (e.g., true/false), their overall test performance will not be so seriously threatened. Other instructors prefer to use only one or two types of questions (e.g., multiple-choice and matching). Because each type of question has advantages and limitations peculiar to it, educators must take into account the purposes of the test, the types of information to be measured, and different learning styles of students.

Test length is another important planning consideration. Educators sometimes attempt to create difficult tests by providing more questions than students can reasonably answer in the time allowed. This does not allow for an accurate evaluation of student performance, and it seriously compromises the validity of the test. Test length should be based upon the maturity of the students and the purpose of the test. Tests that are too long usually test the students' ability to persevere rather than their knowledge of information. Longer tests may favor students who read more rapidly or comprehend unusually well (2). Such traits may be desirable, but they should not be the ones measured by the exam.

Determination of test length is highly dependent upon one's experiences, both with a given level of students and with the materials being considered. Even with experience, significant variations can occur. Different classes taking the same course can vary in their response to the same test questions. Regardless, examinations should not be speed tests in which student performance is affected by rates of item completion, unless, of course, speed is a component of skill mastery in some way. Generally speaking, 90% to 95% of the students should be able to complete a test within the allotted time period (2). Furthermore, it is advisable to make special accommodations for the 10% who may be unable to finish (1).

Tests should be characterized by content validity. That is, the test should reflect student mastery of course content. The use of educational objectives can help to ensure that content validity has been achieved. The test should be constructed with a proportional number of questions reflecting the behavioral objectives for the course or the degree of emphasis placed by the instructor on various course topics. Unless they plan in this way, teachers may inadvertently overemphasize or underemphasize areas within the course. This has

the effect of penalizing students who have placed equal emphasis on all topics included in the course or have misinterpreted educator emphasis upon certain course content (2).

To help themselves balance their tests relative to course content, educators can develop a table of specifications (also called a test blueprint). The primary purpose of a table of specifications is to help the educator balance the test. A balanced test places emphasis on test information in proportions similar to actual classroom activities or course materials (1, 4). Ideally, if 15% of class time (or course materials) is devoted to a particular topic, then approximately 15% of the test questions should address this topic.

Table 2 provides an example of a table of specifications and tells how it can be constructed. Obviously, courses may have many more topics than indicated in this example, but the process is the same. One can also use measures other than class time to derive relative proportions. For example, the relative number of pages of course material devoted to particular topics can be used to indicate proportional numbers of test questions devoted to each topic. Teachers may want to assign weights to specific topics based upon perceived importance, or they may want to balance the questions based upon cognitive learning criteria, such as knowledge, comprehension, and application.

TABLE 2. Table of Specifications

Topic	Class Time Spent on Topic	% of Class Time Spent on Topic	# of Test Questions Devoted to Topic
A	3 hrs.	20%	10
B	2 hrs.	13%	7
C	1 hr.	7%	3
D	3 hrs.	20%	10
E	4 hrs.	27%	13
F	2 hrs.	13%	7
TOTAL	15 hrs.	100%	50

1. In the above example, the teacher decided to give a test with 50 questions.
2. The amount of time spent on each topic was determined and listed. Because the total class time spent on these topics was 15 hours, the relative percentages were determined by dividing the amount of time spent on teach topic by 15.
3. The relative number of test questions was determined by multiplying the percentage of class time spent on each topic by 50.
4. The relative amount of course materials (or other parameters) could be used instead of class time to determine proportionate number of questions to include on the test for each topic.

There are actually a number of advantages to developing a table of specifications. First, it serves to remind the educator of the proportionate amount of time or emphasis being placed on various topics (2). It also helps students to plan their study more effectively. Finally — and, perhaps, most importantly — it can help educators to prepare tests that are valid and appropriate (4). There is a tendency to write test questions on materials for which it is most easy to develop questions. Such unbalanced tests often emphasize knowledge of isolated facts and ignore higher-level cognitive skills, such as understanding, application, and interpretation. This can encourage students to concentrate their studies on materials that are more "testable" (e.g., names, dates, phone numbers, chemical structures, chemical names) as opposed to materials of true learning value (1). The table of specifications results in an analysis of course content and can help to ensure that at least the basic concepts are covered on the test (1). While it takes some time initially to develop the table, it usually saves time in the long run and proves to be an effective means for developing good, functional tests (1, 4).

It is important to write test questions that are clear and concise. Long, involved test questions that lack clarity can cause students to answer incorrectly even when they actually know the correct answer (1). Improperly written test questions reduce test validity and reliability and often place more emphasis on comprehension of test questions rather than on knowledge of course content. The dilemma here often involves the literal meaning versus the implied meaning of words and phrases. Problems with language precision possibly occur more frequently in testing than anywhere else in academic endeavors. Test questions and statements should be analyzed carefully and altered when necessary so that their meaning is unequivocal. This objective of test writing is sometimes impossible to achieve but is, nonetheless, worth educator efforts (1).

Similarly, trick questions should be avoided (1). If students miss test questions because they have been misled by the questions, the test scores will not accurately reflect student knowledge (2). Students recognize questions that have tricked them and, naturally, resent this practice. It gives the impression that the teacher is unfair, and, in fact, this is not a fair practice. It compromises the teacher-student relationship, threatens test validity, and detracts from learning efforts and the inherent value of education. To help students,

instructors should underline, capitalize, or italicize key words. This is especially true for important cues, such as *not* or *least likely* (1).

Teachers should avoid copying material for test questions directly from the text or course materials. Such a practice does not require students to think (2). Rather, they are only required to memorize information and regurgitate it on a test. Memorization is a low-level cognitive skill. Because some students have an innate capacity for memorization, they may perform well on a test even though the test is actually a poor measure of cognitive accomplishment. Many students who perform better on higher-level cognitive tests (e.g., essay) may perform poorly when only memorization skills are actually being measured.

Tests should be prepared several days in advance. This allows the teacher time to review the test, to correct errors before it is given, and to have it printed (2). Advance preparation also allows the instructor to put the test aside for several days, after which the teacher can review it more objectively to see if questions and directions are written clearly and to ensure that questions have only one logical answer (1). Instructors may also want to share the test with colleagues for review. Other teachers will sometimes note ambiguities or difficulties that the author will have missed because of his or her mind-set in writing and reviewing the exam. Because we tend to correct our own mistakes as we review, we frequently fail to spot errors in our own design (1).

It is helpful to write each test question on an index card (2). The cards can be kept in a file by topic, serving as a question pool for future examinations. Over a period of time, a very useful question pool develops, allowing the instructor to vary questions from year to year and yet maintain test validity (1). Information about student responses can be kept on the back of each card (2). Such an analysis allows the educator to identify "good" and "bad" questions. New questions can be added to the file as time permits, and old or inappropriate questions can be deleted. When the time comes to prepare tests, questions can be easily selected, sorted, and arranged for typing.

It is important to critically review test questions periodically. Defective items should be rewritten or replaced with more appropriate items. Sometimes test questions can be improved by merely changing a word or two. Other times, it is much easier to write a new

question than to revise an old one (2). In either case, periodic review helps to ensure that tests are dynamic learning instruments and that they serve legitimate educational objectives. Otherwise, the test and the classroom experience can become stagnated with outdated or inappropriate information.

Instructors should write tests to reflect the course rather than teach the course to reflect tests. Because of the many demands placed upon faculty time, it is tempting to reuse tests. In time, however, teachers may find themselves forcing the course information to conform to the information covered by old exams rather than taking the time and making the effort to update both the course and the exams. In reality, the test should be a reflection of course content and desirable learning objectives. This requires planning on the part of educators so that ample time exists to periodically review tests and test questions. The index card method mentioned above is a particularly convenient approach because it allows for review of individual test questions as time permits.

TEST ORGANIZATION

As with test questions, test directions should be written in a clear, concise manner (2) (Table 3). All instructions should be made as simple and as easily understood as possible. Unfortunately, there are instructors who actually hide some instructions just to see if students will find them. This is an inappropriate educational practice that should be discouraged. The objective of a test is to measure student mastery of course content, not to measure the students' ability to take a test. Trick instructions, like trick questions, serve no

TABLE 3. Points in Test Organization

1. Test directions should be written in a clear, concise, and easily understood manner.
2. Test questions should be grouped to facilitate student performance on and teacher evaluation of the test. Types of grouping include:

    a. Question type (e.g., true/false, multiple-choice)
    b. Topic
    c. Difficulty (proceed from least difficult to most difficult)

3. Questions should be numbered consecutively throughout the test.
4. Questions should be stated completely on a given page.

useful purpose in health science programs. Almost invariably, if test directions are not kept simple and clear, student performance suffers. Test directions should help the students know exactly what they are to do. If directions become too complicated, students may waste valuable time deciphering instructions and acting upon them (1).

Most writers of testing techniques suggest that the testing effort is enhanced when questions are grouped. Grouping is advantageous not only for students but also for teachers because it allows for convenient analysis of groups of questions. There are a number of ways by which questions can be grouped. For example, questions of the same type can be grouped together (e.g., all true/false together, all multiple-choice together) (2, 4). Because each type of question can require a different thought process on the part of the students, grouping by type allows them to concentrate more effectively on content without being distracted by format. This approach also simplifies and minimizes the directions that must be provided on a test (4). A recommended sequence of types in order of increasing difficulty and length of time for a response is listed in Table 4. Tests may use several types of questions, or they may use only one or a few. If more than one type is being used, this organizational hierarchy is considered best.

It is also helpful to students and teachers if questions are grouped according to the topic covered. Again, this affords the students a greater opportunity to concentrate on a particular topic. And, it allows teachers to conveniently analyze student comprehension of various topics within the course. Jumbled topics tend to be more disconcerting to the test-taking effort and can even result in confusion over appropriate answers. Grouping by topic or concept is one of the most useful methods of organizing material on a test (1).

Questions should be placed in order of increasing difficulty

TABLE 4. Recommended Sequence of Test Question Types for Exams (In Order of Increasing Difficulty and Length of Time Required for a Response) (2)

1. True/false
2. Completion (short answer)
3. Multiple-choice
4. Matching
5. Computation
6. Essay

(1, 2). Placing questions in this order helps to bolster student morale at the beginning of the exam and serves to reduce test anxiety. Tensions are eased, and self-confidence is strengthened. This approach also helps to ensure that students can finish the exam in the allotted time because students pace themselves better with this format. Because this can affect student performance, it necessarily affects test validity. When this approach is not used, students may spend inordinate amounts of time on more demanding questions that appear earlier in the exam. Thus, they may not have an opportunity to respond to questions to which they actually know the answers because those questions appear later in the exam. Studies have actually shown that arranging questions from easy to difficult will result in higher scores than arranging items in other sequences (e.g., from difficult to easy, random placement, or placement of easy items among more difficult ones). The advantage of the easy-to-difficult sequence is most pronounced when time is restricted and is much less important when time limits are generous (3).

It is often impossible to arrange questions in order of types, topics, and difficulty all within the same test. At best, educators must strive for a compromise that seems to best facilitate the testing effort (4). In deciding upon a compromise between possible arrangements, it is important to reduce as much as possible those situations that are most likely to prove disconcerting to students. Possibly the best hierarchy of arrangements to follow is question type, topic type, and then difficulty. For example, if completion and multiple-choice questions are to be used, all completion questions should be presented first (in accordance with Table 4). Within this category, questions can be grouped according to topic, and topics can be ordered in terms of increasing difficulty. Educators may also feel that difficulty is really not a significant issue. This may be the case when the various topics are comprised of concepts of varying difficulty or if difficulty is considered to be reasonably uniform throughout the topics covered. In such cases, it may be appropriate to use only question type and topic organization within the test. With many tests it may be convenient to eliminate some organizational categories. For example, if the entire test will consist of multiple-choice questions, then question type is not an organizational concern.

In an effort to reduce the opportunities for academic dishonesty,

some teachers like to use different forms of the same test. The simplest means for doing this is to use a different question order. Obviously, if different questions are used, validity and reliability become greater concerns. And because the students are taking different tests, it is very difficult to guarantee that one version is not actually more difficult than another. If the same questions are used but the order is changed for the different forms, one must be careful that the question arrangements do not affect student performance. From the previous discussion regarding test organization, it becomes apparent that question arrangements *can* actually affect outcomes. Care must be taken to ensure that alternate question arrangements provide similar opportunities for students to perform well and to successfully complete the exam.

Questions should always be numbered consecutively throughout the test (2). Do not start a new numbering sequence with each section of the test. Thus, if it is necessary to refer to a particular question during the test, there will be only one question with that number.

Questions should also be stated completely on a given page (1, 2). It is frustrating, potentially confusing, and time consuming for students to have to flip back and forth between pages in an effort to locate answers or to see the complete question. This can be a particular problem with matching or multiple-choice questions where part of the question is on one page, and the remainder is on another (2). For matching questions, if there is too much information for one page, the materials should be divided into more than one matching section. Then each section can be stated completely on one page. If a multiple-choice question cannot be stated completely on one page, the entire question should be moved to the next page.

INFORMATION TO INCLUDE
ON THE TOP OF AN EXAM

This topic tends to be largely ignored, yet it can be extremely important in helping students to quickly orient themselves to the test-taking effort. Table 5 summarizes information that should be considered. While some of the information is optional, all of these points should be considered relative to a particular group of students, and appropriate information should be included. Course in-

TABLE 5. Information Included on Top of Exam

Course Information
    Name of course
    Course number
    Semester and year
    Time the class meets*
    Name of instructor*

Student Information
    Student name
    Social security number*

Test Information
    Test identification information*
        Which exam (first, second, mid-term, final, etc.)
        Test form (Form A, Form B, etc.)
    Test directions
    Point values of questions (if they vary from question to question)
    Time allowed for taking test*

*Optional information

formation may prove more beneficial to the instructor than to the student. This is especially true if the instructor is teaching more than one section of a course, using similar examinations in more than one course, or saving examinations for a year or more.

It is always a good idea to provide a place for students to write their names on the exam (1). This serves as a reminder to students to identify themselves. If additional information is required (e.g., social security number), a blank for this should be provided as well. Invariably, some students will forget to identify themselves if some provision for identification is not made. This omission should not reflect badly on students. Exam apprehension or excitement can sometimes cause them to forget information that instructors consider obvious.

Test directions are critical for ensuring a valid testing effort. Each section of the test should have its own directions. As stated previously, these directions should be clear, simple, and free of tricky wording. It may also be helpful to visually separate the directions from the questions. This can be accomplished by using different typefaces, using different sizes of letters, placing the directions in a box before each section, or using bold print. Such practices alert the students to the fact that these are sections to which they should direct their attention for instructions. It is also helpful to

include the point value of questions, especially if the point values are not the same for all questions. This allows students to better judge the amount of time they spend on various questions.

Finally, it is advisable to let students know how long they will have to complete the test. However, it is not absolutely necessary to actually print this information on a test. Rather, the instructor should inform the students in preceding class periods how long they will have to take the test because this also helps them to organize their study efforts. The instructor can then remind students when the test is handed out that a certain time period is allowed. Some instructors write this information on the board on the day of the exam. Periodically throughout the exam, instructors may announce the amount of remaining time. Again, this is done in an effort to help students organize their test-taking efforts. This latter practice must be handled with care because the announcement of remaining test time can be distracting.

Usually it is not necessary to provide a title page on exams (a separate page on the front that contains only identifying information). While this page gives a test a more polished appearance, it serves no useful purpose. If teachers are concerned about keeping the front of the test covered, tests can be passed out face down and turned over at the appropriate time.

With appropriate planning and organization, examinations can be challenging learning experiences that provide extremely useful information to all concerned. Without these activities, tests not only fail to achieve the instructor's objectives but also detract from the overall learning effort and erode positive teacher-student relationships in educational environments.

REFERENCES

1. Bergman J. Understanding educational measurement and evaluation. Boston: Houghton Mifflin Co., 1981.

2. Mitchell RJ, Westerman JE. Evaluation in the classroom. Dubuque, IA: Kendall/Hunt Publishing Co., 1972.

3. Sax G. Principles of educational measurement and evaluation. Belmont, CA: Wadsworth Publishing Co., 1974.

4. Ahmann JS, Glock MD. Measuring and evaluating educational achievement. Boston: Allyn and Bacon, Inc., 1971.