THE WILDERNESS INSTRUCTOR

Writing Good Test Questions

Steve Donelan

In a wilderness first aid or first responder course, the best way to test students' skills is to put them through simulated accidents that come as close to reality as possible. However, if we tested all of their skills (both mental and mechanical) in this way, the test would be as long as the course. So we usually need to give written as well as practical exams in order to evaluate student learning.

Traditionally, instructors are supposed to test both knowledge and skills, and there is a tendency to think of written exams as testing only knowledge. Test designers, however, distinguish between questions that test memory (of a fact or a technical term) and questions that test reasoning. And reasoning is a skill, or actually a set of skills, that we teach in our classes. As a result, well-designed questions on written exams can test emergency care skills that may not be adequately tested, or may not be easy to evaluate, in practical exams.

A good analogy is the concept of a "thought experiment" in modern physics. Perhaps the most famous example is the falling elevator used by Einstein to introduce general relativity.¹ He asks the reader to imagine an elevator falling freely from a great height. Then, he poses the question of how an inside and an outside observer would describe events happening in the elevator, such as objects dropped by the occupants, or a beam of light passing through holes in the elevator walls, from side to side.

Similarly, in our test questions, we should put students into mental scenarios that require them to figure out what is going on and what they would do. Then, we will be testing their skills rather than just memory, especially their problem-solving skills. Usually, the point of view implied in these questions is that of the rescuer, who is outside the emergency but needs to find out what things are like from the point of view of the patient. We can make the scenario explicit by describing a situation and then asking questions about it, or we can ask questions that assume an emergency situation.

Scenario questions

In either case, we should begin the process of writing test questions by asking ourselves what we want to test in the students. If we want to test a skill rather than just memory, the next question to ask ourselves is what process we want a student to go through in order to arrive at the correct answer. We should also ask ourselves how well that process corresponds to what students would have to do in a real wilderness emergency. Let's look at an example of a scenario question.

FRACTURES

On a summer mountaineering trip just above the snow line, 2 days from help, one of your 3 companions tumbles down the slope and complains that his lower right leg really hurts. He doesn't feel able to get up and put weight on it. The day is clear and warm. Describe:

- The main steps of a complete patient assessment to rule out other problems.
- How you would examine the leg that is hurting and what you would check.
- What you would do if the leg was crooked and there was no pulse or sensation in the foot.
- How you would splint the leg if it were fractured, using what you carry or would find on the scene.
- How you would get your friend out (just the main points of your plan).

In this example, the student is asked to write short essays rather than select from lists of options in multiple-choice questions. However, the bullet points specify the topics on which students need to write. This helps to guide students through the process of visualizing what they would do for that type of emergency, and going through that process reinforces the question's function as a learning tool. The question format also makes it easier for the instructor to grade the answer, since answers will generally be better focused and organized than with an open-ended question, and each of the 5 bullet points counts for one fifth of the total score.

Direct correspondence to PO Box 1227, Berkeley, CA 94701 (e-mail: donelan@mindspring.com).

Nevertheless, essay questions require much more time and thought to grade than multiple-choice questions and can be difficult to decipher if the answers are handwritten. An alternative is to provide 4 or 5 multiple-choice questions based on the scenarios. For example, you could give the student 4 options for the answer to the first bullet point, only 1 of which listed the correct steps of a patient assessment in the correct sequence.

In either case, what are these scenario questions testing? The **topic** is fractures in a wilderness situation. The **skills tested** are the student's ability to:

- Recognize and locate a fracture, and find or rule out other problems.
- Check distal functions.
- Align an angulated limb.
- Improvise a splint.
- Make an evacuation plan.

Multiple-choice questions

Could we test some of these skills with independent questions that do not refer to a common scenario? Let's take the third bullet point as an example and write a multiple-choice question on it.

In a wilderness accident, if your patient has an angulated fracture with impaired distal functions, you would usually:

- (a) Splint the limb as it lies without trying to align it.
- (b) Manipulate the limb and try to bring the broken bone ends together.
- (c) Apply gentle axial traction to align the limb, and then check distal functions again.
- (d) Not attempt to splint the limb since you could cause further damage.

The **topic** of this question is wilderness fracture treatment. Let's look at what information the statement (or **stem**) gives to students.

- The accident is in a wilderness, not an urban, setting.
- The patient has a fracture.
- The fracture is angulated.
- Distal functions are impaired.

In addition to providing the information that students need to select the correct answer, the stem also reinforces several important points:

- Treatment in a wilderness situation may differ from treatment in an urban situation.
- Angulation is an obvious sign of a fracture.
- You should always check distal functions in a fractured limb.

Being reminded of the third point should trigger a reasoning process in students:

- Why is it important to check distal functions?
- What would happen to the limb if distal functions were impaired?
- What could I do about it?

This reasoning process should lead students to the correct option, (c).

All of the other choices (the **distracters**) are plausible, however. So students who were not paying attention in class, or who didn't think the question through, could easily select an incorrect option. For the instructor, therefore, this question can help identify students who might not give the appropriate treatment for a fracture in the wilderness and might need further training or review.

You could put the same question in several different formats. For example, you could make it a statement and ask if it was true or false, which in effect reduces the options to 2. If you completed the statement with either option (a) or option (c), you would be strongly emphasizing the difference between wilderness and urban fracture treatment. You could also use an illustration of an angulated fracture to present information about the problem. Or you could break up the options so that the correct answer would be a combination of several. For example, if "axial traction" was option (b), and "checking distal functions again" was option (c), then the correct choice would be "(b) and (c)." Another way to present the options is by requiring students to match items in 2 columns (eg, different types of injuries with treatment).

Analyzing test questions

How could we tell whether the question was effective? There are several ways to analyze the results of a multiple-choice test. First, we can estimate the difficulty of the question by finding the percentage of students who selected the correct answer. If that percentage is very small (less than 40%), then there are several possibilities:

- The question really was too difficult for the class and the level of training you provided.
- The question was poorly written (obscure or ambiguous).
- You didn't teach this particular topic very well.
- Many of the students missed the relevant class.

If the percentage is very high (over 90%), then you should also consider several possibilities:

- The question was too easy.
- The question gave away the answer.

- You did a great job of teaching the topic.
- You have a very good group of students.

Another way to analyze the question is to correlate the student responses to the question with their total score, and then compute the average score of those who selected the correct option. If their average score is above the class average, then you have a positive correlation between the choice of the correct option and the overall performance of students on the test, which validates the question. If the average score of those who selected that option is lower than the class average, however, then you need to ask yourself why so many of the best students disagreed with you and why so many of the poorer students agreed with you. Was your choice of the correct answer really wrong or arguable? Did it contradict the textbook or what you taught in class? Or was the question just so poorly written that a careful reading would point the better students toward another option?

A poorly written question can give away the correct answer with grammatical clues. Ask yourself, for example, what is wrong with the following question: The best backcountry treatment for a midshaft femur fracture is a:

- (a) Soft splint.
- (b) Traction splint.
- (c) Improvised fixation splint.
- (d) Air splint.

The use of the article “a” in the stem rules out options (c) and (d), which require “an.” The grammatical clue can be removed by replacing “a” with “a(n).” The following question gives another example of a grammatical clue: The most likely causes of wound infection in the wilderness are:

- (a) Inadequate bandaging.
- (b) Delayed evacuation.
- (c) Wound contamination and inadequate cleaning.
- (d) Failure to apply antibiotic ointment.

The use of the plural verb “are” in the stem rules out options (a), (b), and (d). The clue can be removed by replacing “causes” with “cause(s)” and “are” with “is/are.” Alternatively, you could make all the distracters plural.

You can also make a question too easy by making the correct answer too obvious or the distracters too implausible. For example, what is wrong with the following question?

If a winter camper has deeply frostbitten feet, you should:

- (a) Rub the feet with snow.

- (b) Slap sensation back into the feet.
- (c) Thaw the feet in hot water and continue the trip.
- (d) Evacuate him or her to a hospital.

While some students who did not do the reading or pay attention to the lesson may choose option (a) or (c), nobody is likely to take option (b) seriously. However, the strongest clue is that option (d) is so obviously the safe choice that it would be hard to disagree with it. So students who are adept at test taking will immediately rule out the first 3 options. You could improve the question by making some of the distracters more plausible and making the correct choice more like the distracters (eg, “(d) Have the patient walk out on the frozen feet if necessary.”).

Another clue that enables sophisticated students to eliminate options is the use of absolutes such as “always” and “never.” Any statement about emergency care that includes one of these words is likely to be false, or at least very arguable. Words expressing probability, such as “usually” or “seldom,” will make much more plausible options.

Bear in mind, however, that statistics do not always tell you what is going on. Students may do poorly on written tests for reasons that have nothing to do with their knowledge or skills. For example, they may have a reading problem, a language problem, or a test-taking problem. Classes in the United States are often culturally diverse, with students for whom English is a second or third language. Even native speakers may have problems reading and understanding written questions. English has a higher percentage of dyslexia than any other language, because it is so heterogeneous, and its grammar and spelling are so inconsistent. Even students with advanced degrees may be dyslexic. So if some students are doing well in the practice sessions and scenarios but having trouble with the written tests, you need to find out why and develop alternative ways to test them. These alternatives might include giving tests orally, finding translators, or having the tests translated into other languages. If they are in multiple-choice format, grading them will not be a problem.

Selecting topics to test

How do you decide what topics to include in your written exams and what skills (related to each topic) you need to test? Go back to your course plan for the topics to be tested.² Then, go back to your objectives for each lesson plan.^{3,4} What did you want students to be able to do at the end of each lesson? Verbs expressing lesson objectives include:

- Recognize;

- List;
- Describe;
- Distinguish;
- Explain; and
- Demonstrate.

Whether you use an explicit scenario as the basis for several questions or imply a scenario in the question, your goal should be to put the students through a reasoning or decision-making process that corresponds to, and prepares them for, what they would do in a real wilderness emergency. If the question tests a skill that would be irrelevant or unnecessary in any conceivable wilderness emergency, then it probably does not belong on the exam. Questions that do not put students through any process except remembering a term or a fact are also of little use in evaluating or predicting their performance.

How do you decide the number of questions for each topic? Again, go back to your course and lesson plans. How much time did you actually spend teaching each topic? How many skills were involved? How complex were they? How important were the problems that the skills were designed to manage? Remember that the importance of a problem is measured both by how common it is and how serious it is. Students need to be adept at managing common problems, such as cleaning and bandaging open wounds. But they also need to be able to recognize and manage uncommon problems that can kill or disable a patient, because their memory of these problems and what to do about them will not be reinforced by frequent practice. For example, few people will ever encounter a sucking chest wound in the wilderness. But if they do, they need to respond immediately with the correct treatment. So one function of a written exam should be to reinforce recognition of and response to serious problems that occur infrequently. And a little humor is not out of place if it helps to reinforce an important point. For example: True or False: Do not cover a sucking chest wound—it may be the victim's only way to breathe.

What is a skill?

When you are writing test questions on a topic, visualize what students would have to do in a real situation involving that topic. The skills they will need to employ may include:

- Observation (of the scene and the patient);
- Analysis (figuring out what the problems are);
- Planning patient care and setting priorities;
- Organizing patient care and rescue;
- Organizing and improvising equipment;

- Communicating with patients, rescuers, and bystanders;
- Writing reports;
- Providing leadership and teamwork;
- Coping with psychological stress in oneself and others; and
- Dealing with the mechanics of emergency care (bandaging, splinting, etc).

Notice that most of these skills are done with the brain rather than the hands, and written exams are especially good at evaluating and reinforcing these nonmechanical skills. Ask yourself which skills the following questions test and reinforce. Then, see how many similar (or better) questions you can write yourself on wilderness emergency care topics.

A good test for high-altitude cerebral edema is to:

- Have the patient count the number of fingers that you hold up.
- Check whether the pulse is abnormally rapid.
- Have the patient try to walk a straight line, heel against toe.
- Check the chest for crackly breathing sounds.

After a wilderness accident, 1 member of the party is loudly assigning blame, and several others are arguing with him. You should:

- Correct him and settle the argument.
- Ask him to help care for one of the injured people.
- Ignore him while you take care of the injured.
- Bring everyone into the discussion and settle the point democratically.

The leader of a backboarding rescue forgets to do a complete patient assessment. You should:

- Ask her whether she would like you to check vital signs and get a SAMPLE history.
- Challenge her qualifications to lead the rescue and offer to take her place.
- Remain silent since it is more important that the rest of the rescue go smoothly.
- Correct her mistake in front of the group so that she will be more careful after that.

True or False: The most reliable sign of heat stroke is red, hot, and dry skin.

In a wilderness accident with many casualties, which patients should receive care first?

- Those who are complaining the loudest.
- Those who are breathing but have altered vital signs.
- Those who do not start breathing and have no pulse.
- Those with obviously severe injuries such as fractures.

Conclusion

Good written exams should evaluate and reinforce emergency care skills. You can identify these skills by going back to the objectives in your lesson plans. As you write a test question, ask yourself what process you want the students to go through in order to arrive at the correct answer and how well the process prepares them for what they would do in a real wilderness emergency. Written exams can easily test the nonmechanical skills. But it is also possible to test some aspects of mechanical skills by writing questions that require students to visualize and mentally walk through the skills. Finally, remember that evaluation and grading are not the only functions of written exams. They should also serve as powerful

learning tools that give students a review of the whole course as well as a mental rehearsal of the skills they will need to cope with real emergencies.

References

1. Einstein A, Infeld L. *The Evolution of Physics*. New York, NY: Simon and Schuster; 1938.
2. Donelan S. Teaching wilderness emergency care. *Wilderness Environ Med*. 1999;10:40–43.
3. Donelan S. Teaching emergency care skills. *Wilderness Environ Med*. 1999;10:125–127.
4. Donelan S. Secrets of a successful lecture. *Wilderness Environ Med*. 1999;10:185–188.