

IMPROVING YOUR WRITTEN EXAMS

Debra Cason, Program Director and Associate Professor, Emergency Medicine Education
University of Texas Southwestern Medical Center at Dallas

November, 2008

Debra.Cason@utsouthwestern.edu

Objectives: Upon completion of this class the learner will be able to:

1. Define validity and reliability
2. Discuss factors affecting validity and reliability
3. List and describe 7 principles of written exam improvement
4. Apply principles of item and test construction to an exam

EXAMS are “powerful educational tools that serve at least four functions: 1) tests help you evaluate students and assess **whether they are learning** what you are expecting them to learn, 2) well designed tests serve to **motivate and help students structure their academic** efforts, 3) tests can **help you understand how successfully you are presenting** the material, and 4) tests can **reinforce learning** by providing students with indicators of what topics or skills they have not yet mastered.” Barbara Davis, 2001, *Tools for Teaching*

- ◆ When should you use written exams?
- ◆ When should you use objective multiple choice exams?
- ◆ Who is responsible for writing written multiple choice exams?
- ◆ Who should understand and practice principles of exam writing and evaluation?

I. VALIDITY = relevance

Does the test measure what it is supposed to?

The student with more knowledge in the content area should score higher than a student that does not know as much.

Types of validity:

- a. Content validity—follows course objectives and content
Tie each exam item to an objective and make sure a representative “biopsy” of the content is included in the test.
- b. Face validity—common sense validity.
Have a colleague look at the item and agree that the item “makes sense”.

- c. Construct validity—construction follows certain principles
- d. Criterion related validity—how does this test compare with other external measures? Does the student perform well?

II. RELIABILITY = consistency

Is the test reproducible? How precise and accurate is the test? Will the student score the same on other tests of same content?

Does the test measure true factors (such as knowledge about a particular subject) or does it measure chance?

III. PRINCIPLES of EXAM IMPROVEMENT

1. Be clear of the purpose of each exam
2. Base each test item on an objective
3. Blueprint the exam to appropriately cover and weigh topics
4. Construct items appropriately
5. Make adaptations from published item banks
6. Obtain multiple sources of “expert” item review (include medical director).
7. Use “higher cognitive” items vs. recall items

Cognitive Domain Levels

1. Knowledge—recall, comprehend (define, discuss, describe, list...)
2. Application—apply, use (predict, apply, compute...)
3. Problem solving—analyze, synthesize, evaluate (evaluate, judge, differentiate...)

8. Write clear, complete directions for each test or each section of the test

9. “Prepare” examinees for test
 - the “sleep and no meds” lecture
 - similar type of exam items previously
 - clear objectives
 - subjects that will be covered
 - discuss rules

10. Pay attention to exam security—before, during and after the exam
 - appropriate spacing
 - use cover sheets
 - stroll the classroom
 - post exam review
 - don’t give exams back to keep
 - avoid teaching the test
 - continually revise high stakes exams

11. Evaluate your exam and items after it is administered.
 - student feedback on items
 - difficulty index
 - did the high achievers do better than the low achievers?

**Sample Blueprint
Cardiac Disorders**

| Category | cognitive # objectives | # items |
|--------------------------------------|-----------------------------------|----------------|
| 1. Atherosclerosis | 3 | 1 |
| 2. Angina | 8 | 4 |
| 3. Myocardial infarction | 10 | 6 |
| 4. Heart failure | 4 | 2 |
| 5. Cardiac arrest | 8 | 6 |
| 6. Aortic aneurysm | 2 | 0-1 |
| 7. Arterial occlusion | 2 | 0-1 |
| 8. Pulmonary embolism | 3 | 0-1 |
| 9. Hypertension | 4 | 1 |
| 10. Other (pericarditis, etc) | 4 | 1 |

**Sample Blueprint
EMT FINAL EXAM**

| | |
|-------------------------------|-------------------|
| 1. Role of EMT | 5-7 |
| 2. Airway | 14-16 |
| 3. Patient assessment | 7-8 |
| 4. Medical Emergencies | 16-18 |
| 5. Trauma | 12-14 |
| 6. Special populations | 10-12 |
| <u>7. Operations</u> | <u>6-7</u> |
| TOTAL | 70-82 |

Exam Item Construction

Principles relating to the Item as a Whole

1. There should be only one clearly correct or best answer.
2. Unless reading ability or verbal skills are being measured, the reading difficulty of items should be low in relation to the group taking the test.
3. The items should be original, not sentences lifted directly from the text.
4. Items based on opinion or authority should state whose opinion or authority provides the basis for the items.
5. Trick questions should be avoided.
6. Items should be useful, worthwhile in the content area being tested, not trivial.
7. The meaning of items should be clear.
8. Items should be constructed at a level of difficulty appropriate to the examinees.

Principles relating to the Stem

1. The stem should have a single thought which presents the problem clearly.
2. The stem should contain as much of the item as possible.
3. Negative statements should be avoided if possible; if negative words are used, attention should be called to their negative aspect by underlining or capitalizing. Avoid double negatives.

Principles Relating to Options

1. Options should be as homogeneous as possible.
2. Options should be arranged in some logical order, if possible.
3. The wording of the correct response should not provide a “sound alike” or direct work clues with the stem.
4. Since indefinite terms such as “generally”, “frequently”, “for the most part” tend to be in true options, and absolute terms such as “always” and “never” tend to be in false options, options should not provide these clues. (The stem of a multiple choice item may include such terms as needed.)
5. Options should be grammatically correct in terms of the stem.
6. The correct answer should provide not irrelevant grammatical clues.
7. Options should be similar in content (homogeneous). This type of options make items more difficult; options which are dissimilar (heterogenous) make items easier.
8. All items should be plausible.

TEST

Directions: For each item below, circle the one best answer.

1. Which of the following pairs has won the greatest number of Abby awards?
 - a. Jones and Smith
 - b. Smith and Taylor
 - c. Smith and White
 - d. White and Allen

2. How many pounds of pressure are exerted by a callam?
 - a. 2.6
 - b. 150
 - c. 260
 - d. 2600

3. The stanon is aided by a
 - a. anstel
 - b. immon
 - c. octal
 - d. port

4. The stanon frequently overheats because
 - a. all grestels are bilious
 - b. no immon are directly fectitious
 - c. octals are usually casable
 - d. ports are always actial

5. Stamation normally occurs when the
 - a. ansetls rupture
 - b. immon falls and the denton is in place
 - c. octal rotates easily
 - d. ports pass the carm