



Center for Teaching Excellence
Hampton University
Teaching Matters

March/April 2016

Volume 11, Number 2

Featured Article:

An Introduction to Test Planning, Creating Test Items, Conducting Test Item Analysis, Part I

By: Dr. Simone Heyliger, Associate Professor, Pharmaceutical Sciences and

Ms. Deborah Hudson, Assistant Professor, Pharmacy Practice

The creation of high quality test items is an essential skill for all college instructors. However, many professors have little formal training or experience in the development and analysis of the test questions. Concerns are further compounded by new accreditation standards which require individual academic units or institutions to provide evidence that exams are valid and actually test students on the competencies outlined in syllabi and accreditation standards. The goal of this article is to provide basic steps regarding the development of test items as well as the disadvantages and advantages associated with certain item formats. In addition, the article will also provide information on how to utilize test item statistics to revise exams and enhance exam questions.

STEP 1: Creating Measurable Learning Objectives

The development of a good test question starts with the creation of student intended learning objectives (SILOS). A learning objective is defined as a statement of outcome that describes the skill or information that the learner (student) will acquire after engaging in a learning activity (instruction). Learning objectives are important because they: 1) focus students on what is important, 2) facilitate lecture planning and test assessment development and finally, 3) because they are used increasingly to construct Curricular Maps and to perform Curriculum review and analyses. Typical learning objectives are based on learning models such as the Revised Bloom's Taxonomy (Krathwohl, 2002), and thus will include "Bloom's action verbs" as part of the learning objective (Refer to the Appendix). Many of these action

verbs are arranged according to the rank of cognition (thinking skills) that are required to master given material or specific learning activity. According to the Revised Bloom's Taxonomy, lower order thinking skills are associated with memorization, whereas the highest order thinking skills are associated with prediction and creating. Hence, when developing learning objectives, instructors must remain vigilant to insure that their instructional strategies complement learning objectives and are in total alignment with the levels of difficulty and expectations of a given test question. Table 1 provides examples of the evolution of a learning objective.

| Table 1. The Evolution of a Learning Objective |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Phase 1: <i>Know how to interpret Arterial Blood Gases (ABGs) lab values and to use the compensation formulas</i></p> <p>COMMENT: Student does not understand how “KNOW” will be quantified (assessed/ tested) and thus have no idea of the level of preparation that is needed. Therefore, this learning objective is unmeasurable.</p> |
| <p>Phase 2 (Revised Learning Objective):</p> <ul style="list-style-type: none"> ▪ Use the ABG diagnostic algorithm to identify the primary (most prominent) acid-base disturbance (acidosis or alkalosis) that is present in a patient profile (case study) which contains patient's ABG lab values, symptoms, clinical findings and medical histories. ▪ Apply the compensation formula associated with the primary acidosis or alkalosis condition to determine if the body is attempting to correct the abnormality ▪ Calculate the Anion gap and Delta ratio ▪ Compare the patient's Anion gap and delta ratio values against normal ranges to determine if there are hidden metabolic acidosis and metabolic alkalosis <p>COMMENT: Instructor's expectation was expansive and actually required students to perform different tasks. It is better to have multiple learning objectives which provide stepwise approaches to addressing content and knowledge. Students now have a fairly good understanding how they must approach preparing for the subject matter.</p> |

STEP 2: Test Planning (Test Blue Printing)

Many instructors (myself included) are guilty of constructing exams on short notice because of time constraints. However, in order to create a good exam which adequately assesses students' competencies, significant planning should be performed prior to the construction of an exam. This section provides suggestions on how to plan an exam. By developing test blue prints, instructors will be able to analyze results after the exam is administered to students. This is particularly important when a significant portion of the student population performs poorly on an exam. In addition, test maps can be used as documentable evidence that a particular course or curriculum is addressing competencies/ standards outlined by accrediting bodies. Table 2 outlines the steps that should be undertaken to plan an exam, while figure 2 represents an example of a test planning matrix.

Table 2. Test Planning Process

- A) Determine which Chapters, Topics and most importantly which **SILOs** will be covered on the exam.
- B) Classify **SILOs** based on Bloom’s Taxonomy (List cognition level).
- C) Determine the percentage of test that is dedicated to each **SILO**.
- D) Decide which test writing strategies will be used to assess each **SILO**.
- E) Group and number test items based on **SILOS**.

| SILOS | % of Test | Test Item Strategy | No. Test Q per Item | Bloom Level | Q No assoc. w/ SILO |
|--------|-----------|--------------------|---------------------|----------------------------|---------------------|
| SILO 1 | 50% | Multiple Choice | 15 questions | Level 3 (Applying) | 1-5, 16-20, 26-30 |
| | | T and F | 5 | Level 1 (Remembering) | 6-10 |
| | | Matching | 5 | Level 1 (Remembering) | 11-15 |
| SILO2 | 20% | Calculations | 10 questions | Level 4 (Analyzing) | 21-25, 31-35 |
| SILO3 | 10% | T and F | 5 | Level 1 (Remembering) | 36-40 |
| SILO4 | 10% | Multiple Choice | 5 | Level 2 (Understanding) | 41-45 |
| SILO5 | 10% | Short Answer | 5 | Level 4 (Analyzing) | 46-50 |

Figure 2. Test Planning Matrix

(Adapted from http://orgs.bloomu.edu/tale/documents/OAE4_TestBlueprinting.pdf).

STEP 3: Choosing the Appropriate Test Writing Items

There are numerous types of test questions that can be utilized to test students’ knowledge. The most popular strategies include matching, true and false, multiple choice questions (MCQ), fill in the blanks, short answers and essays. Each one of these item types has advantages and disadvantages. Table 3 summarizes these test item strategies. The remainder of this section will focus on the development of multiple choice exam questions because that is the preferred test writing strategy of standardized entrance and licensure examinations. In addition, multiple choice questions are extremely versatile and can be used to test many cognitive skill levels (as defined by the Revised Bloom’s Taxonomy). Lastly, the rigor and quality of multiple choice questions can be easily quantified using various test analyses and statistical software.

Table 3. TEST ITEM WRITING STRATEGIES

| Test Item | Advantages | Disadvantages | Bloom's Level |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|
| True/False | <ul style="list-style-type: none"> • Easy to create • Students can answer numerous question short time | <ul style="list-style-type: none"> • Does not test knowledge • Students often guess • 50% chance of student answering correctly without any prior knowledge of subject | 1 &2 |
| <p>Comments: This item strategy is not recommended for use because students have a high probability of guessing correctly.</p> | | | |
| Essays | <ul style="list-style-type: none"> • Students must utilize higher level learning cognitive skills • Can assess students grasp of broader content of knowledge as well as relationships between complex concepts • More opportunities for students to express themselves and to display their creativity | <ul style="list-style-type: none"> • Highly subjective • Students more likely to contest grades • Requires the development of a detailed rubric • Time consuming for students to complete • Time consuming for instructors to grades • Limited numbers of questions that can be asked on a test • Students' writings may be unfocused and disorganized | 5 &6 |
| <p>Comments: Essays are time consuming to grade and are not logistically feasible when teaching large classes. One has to be careful of not introducing grading biases. For students with better vocabulary. Moreover as the evaluator becomes tired, he or she tends to award better grades to essays towards the end of the pile.</p> | | | |
| Matching | <ul style="list-style-type: none"> • Useful when covering discipline with lots of facts • Useful for comparing related ideas • Useful for testing cause and effect associations • Easy to score | <ul style="list-style-type: none"> • Time consuming for students • Promotes Student guessing • Does not evaluate knowledge of isolated facts or concepts • Promotes rote memorization | 1 &2 |
| <p>• Comments: Matching questions should be used sparingly because they do not test higher orders of learning and they tend to promote rote memorization.</p> | | | |

Multiple choice exams consist of a “**stem**” which states the problem, the correct answer and several wrong, but plausible, choices (called “**distractors**”). The key to writing good multiple choice exam questions is to avoid the following errors:

- Grammatical cues which provide clues to the appropriate answer (the correct answer is the only phrase that has correct subject/verb agreement.)
- Construction of sentences which provide non-content related clues that indicate the choice that is the correct answer (sentence normally has the most detail or is the longest choice).
- Repetition of words (the answer contains the same words as the stem).
- Utilizing the same choice designation as the correct answer (some writers may favor designating choice “C” as the correct answer).
- Writing long stems with superfluous information.

The following examples are provided to demonstrate the mechanics of developing good multiple choice questions.

Example of a convoluted, verbose Stem:

- *A skin biopsy of a patient who was tested with purified protein derivative (PPD) to determine previous exposure to Mycobacterium tuberculosis develops induration (swelling and redness) at the site 48 hours later would **MOST PROBABLY** reveal_____.*

Rewrite:

A patient is skin-tested with purified protein derivative (PPD) to determine if he was previously exposed to *Mycobacterium tuberculosis*. Forty-eight hours later, he develops an induration (swelling and redness) at the site of the skin test.

A skin biopsy of the swollen site would MOST LIKELY reveal the presence of_____

Verbose information should be simplified or eliminated and then separated from the stem.

Example of Repetition of Words:

- Autoimmune diseases can be characterized by all of the following **EXCEPT**:*
- A) **Autoimmune** disease has a tendency to occur more frequently in women
 - B) There is a genetics (HLA-D) may predispose an individual to develop an **autoimmune** disease
 - C) Autoantibodies are often present in serum of patients who suffer from **autoimmune** disorders
 - D) In **autoimmune** diseases, molecular mimicry by pathogens lead to the production of antibodies which react to host's tissues
 - E) Occurs when B-cells fail to class which from IgM to IgE

The answer to the above question was “E.” Approximately 93% of the students chose the correct answer. Item analysis revealed this was a poor question. Each distractor contained the word “autoimmune”. Thus, astute test takers, using the process of elimination, would automatically gravitate to choice “E” because the correct answer differs significantly from the other choices.

Example of an unfocused stem

- Which of the following statements is Correct?
- A. HIV is transmitted by skin being splashed with blood from an HIV infected individual
 - B. HIV is transmitted by exchanging a kiss with an HIV positive person
 - C. HIV is transmitted by engaging in sexual contact with a male using a condom
 - D. HIV is transmitted by exchanging body fluids percutaneously via needles
 - E. HIV is transmitted after being bitten by a mosquito.

There are several problems with this question. First, the stem is not associated with a learning objective. Students should be able to read a stem and automatically identify the subject area that is covered by the stem. Hence, students are placed in a position where they have to guess the intent of the question.

Rewrite

- Which of the following represents the **greatest risk** for acquiring HIV disease?
- A. Being splashed with blood from an HIV infected individual
 - B. Kissing an HIV positive person
 - C. Engaging in sexual contact with a male using a condom
 - D. Exchanging body fluids percutaneously via needles
 - E. Being bitten by a mosquito

In the rewrite, students who read the question are able to determine that the question is addressing mode of transmission of the virus.

In summary, writing a good test question is essential for testing student achievement. Care should be taken to avoid factors which undermine the validity of the exam. Other practices which should be avoided include using negative phrases such as “**EXCEPT.**” Students often overlook the word and it causes confusion regarding the intent of the question. When crafting distractors, each distractor should be of similar length. A common mistake that is made by test writers is that they tend to make the correct answer the longest sentence. Astute test takers will automatically choose the longest choice. In addition, when including numbers, the distractors should be ranked numerically. Single word distractors should be ranked alphabetically. Do not include distractors that are obviously wrong. A good distractor is a reasonable choice (though clearly incorrect) and thus a small subset of students should select each distractor. Distractors which are not chosen by any students should be revised or eliminated from future exams. Once the exam is administered, the test should be evaluated for reliability and to identify any deficits in instructional strategies. Part II of this article will discuss how to use test and item analysis to

assess student knowledge and to identify curricular holes and poorly designed or even mis-keyed questions.

REFERENCES AND USEFUL LINKS

Haladyna, Thomas M. Developing and validating multiple-choice test items, 2nd edition. Lawrence Erlbaum Associates, 1999.

Krathwohl , D. (2002) A Revision of Bloom's Taxonomy: An Overview in Theory into Practice, 41 (4): 212-218 (http://www.unco.edu/cetl/sir/stating_outcome/documents/Krathwohl.pdf)

Teacher & Educational Development, University of New Mexico School of Medicine “Effective Use of Performance Objectives for Learning”
<http://ccoe.rbhs.rutgers.edu/forms/EffectiveUseofLearningObjectives.pdf>

Vanderbilt Center for Teaching Excellence <https://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions/>

Appendix

Revised Bloom's Taxonomy Action Verbs

