

DEPARTMENT OF THE ARMY
 U.S. Army Medical Department Center and School
 Fort Sam Houston, Texas 78234-6100

AMEDDC&S Pamphlet
 No. 351-17

2 April 1993

Schools
 JOB PERFORMANCE AID: TEST ITEM DEVELOPMENT

TABLE OF CONTENTS

		<u>PARAGRAPH</u>	
CHAPTER 1	Preface	1-1	
CHAPTER 2	Guidelines for Preparing Tests	2-1	
CHAPTER 3	Steps in Writing a Test	3-1	
CHAPTER 4	Guidelines for Writing Test Directions	4-1	
CHAPTER 5	Guidelines for Writing Multiple Choice Test Items	5-1	
CHAPTER 6	Guidelines for Writing Alternative Choice Test Items	6-1	
CHAPTER 7	Guidelines for Writing Matching Test Items	7-1	
CHAPTER 8	Guidelines for Writing Short Answer/Completion/List Test Items	8-1	
CHAPTER 9	Guidelines for Writing Essay Test Items	9-1	
CHAPTER 10	Guidelines for Writing Performance Test Items	10-1	10-1
CHAPTER 11	Guidelines for Formatting a Test	11-1	
CHAPTER 12	Guidelines for Scoring Tests	12-1	

CHAPTER 1

PREFACE

1-1. **Purpose.** This document is a job performance aid for developing and evaluating test items. It is designed to serve as a reference for training developers when preparing tests pertaining to the student assessment program for resident instruction at the U.S. Army Medical Department Center and School (AMEDDC&S). This document should be used in consonance with the AMEDDC&S Regulation 351-19, Testing Procedures and Policies.

1-2. **References.** Related publications are listed in Appendix A.

1-3. **Scope.** This Pamphlet describes the process for developing test items for the AMEDDC&S.

1-4. **Overview.** Testing is used after training to determine whether or to what degree trainees learned what they were intended to learn. How detailed and inclusive a test should be depends on the use for which it is intended. If the need is to place trainees in remedial courses, place them at various points within a course, or exempt them from a course, the test must be sufficiently broad to cover all course prerequisites, entry behavior, and all of the learning objectives. It must be detailed enough to discriminate between those who can perform and those who cannot perform.

a. If the purpose of the test is to give trainees information on how they are progressing, the test should include test items that can be scored quickly to give the student the appropriate feedback. This feedback may also be used to make decisions to consider students for remedial instruction.

b. When a course is being developed, there is an additional use for test results. If students are tested frequently, the scores provide a record of how well the instruction is attaining the prescribed standards. Errors by students may pinpoint problems in instruction/instructional materials. Early achievement of goals or higher than expected standards are indicators that instruction should be reviewed and adjustments made as documented.

1-5. **Applicability.** This job performance aid is designed for use by anyone who is charged with developing tests. It should be used in consonance with regulatory documents and school SOP and policy.

1-6. **Exception to Policy.** Given the diversity of AMEDD courses, certain detailed standards may not be applicable in all cases. Graduate programs (those which grant degrees) may follow the established testing structures and policies and procedures of affiliated colleges and universities. Courses which are required to follow civilian examination and certification programs may follow the testing structures and policies of the affiliated program.

1-7. **General.** Testing is an essential component of training. Testing serves to demonstrate that end-of-course objectives have been achieved. Most learners dislike being tested but do accept testing as a necessary part of training. A major job of the training developer is to design reliable and valid instructional tests that will measure learner performance adequately and not try to trick the student or purposely fail the student. A group of test items cannot necessarily be called a test. After a group of test items has been validated, the training developer in consonance with the subject matter expert (SME) must decide the number of items to be included and the order in which the items are to be presented to the learners. Test directions and a scoring key must also be developed. Each of these steps must be taken carefully if the test items do what they are supposed to do. A good test must measure accurately what it is supposed to measure; measure consistently what it is designed to measure; be fair to the students; differentiate between students who can and cannot accomplish the objectives; contain a liberal sampling of items across all training objectives; and be easy to use. All of these factors are interdependent and affect each other. These factors should be present in a good test.

CHAPTER 2

GUIDELINES FOR PREPARING TESTS

2-1. A valid test item is derived from a learning objective. Test items are developed that are parallel to and measure the learner's ability to achieve what is stated in the objective. The training developer/test writer must ensure that the test item requirements match the objective's requirements. The training developer should ensure that each test item tells the trainer not only what the student knows but also what the student does not know. Some objectives may have to be rewritten so they can be measured.

2-2. Types of Tests.

a. Full Performance Tests: Actual performance of the task using authentic job conditions and standards.

b. Partial Performance Tests (based on interim/predictive objectives): Simulation of the actual task or a change in the task characteristics. For instance, flight simulators are used instead of actual aircraft or actual combat triage situations may be changed to simulated combat triage situations using moulaged patients. Keep in mind that the more a task is changed, the greater the loss of fidelity.

c. Written Tests: Written tests are usually effective in assessing the acquisition of knowledge of facts, principles, and procedures at the various levels of the cognitive domain. Knowledge testing is generally a reasonable substitute for performance testing only when it can be reasonably assumed that anyone possessing the knowledge could perform the task (no additional training is required). However, all performance tasks require a certain amount of knowledge that can be measured effectively with a written test.

2-3. General Rules for Writing Tests. The following guidelines should be followed when writing any type of test.

a. All questions should be based on learning objectives. Well-defined, specific objectives, stated in terms of behavior that can be seen, heard, or felt, form the basis for any evaluation tool. The action, conditions, and standards stated in the objective must be included when preparing the test. Do NOT add anything to the test item that was not included in the objective. A common error is to add sequence to the test item

when sequence is not mentioned in the objective. If a student must perform the behavior in sequence, the sequence requirement must be stated in the objective.

b. Plan ahead. Test construction is time-consuming; good tests take time and should be planned in detail well in advance instead of as an afterthought.

c. Write succinct/understandable test items. The stem and alternatives should be written so that the meaning is clear. Include all possible information in the stem.

(1) When words or phrases are common to all choices, they should be placed in the stem.

(2) The stem should stand alone so that students who know the subject can answer the question without looking at the choices.

(3) Stems should contain only material relevant to the solution. The stem should not flag or give away the answer by irrelevant details, grammatical discrepancies, or other clues. Since the longest answers are usually correct, the alternatives should be approximately the same length.

d. Develop test items that present problems that require specialized knowledge. Avoid test items that can be answered on the basis of intelligence or general knowledge. Students should be required to solve realistic job-related problems. Base test items on realistic, practical down-to-earth, performance-oriented job situations.

e. Avoid verbatim copying from the text. This encourages memorization rather than application. Some things should be learned verbatim, but to over use this could result in the student selecting the correct answer without knowing what it means or how to react under different conditions.

f. Avoid trick test items. To be useful, test items must be clear-cut and without deception. Questions are to be used as an evaluation and teaching device and must not be used to deceive the student. If trick test items are used, the student begins to look for a trick in every test item. This results in the selection of wrong answers by the student while seeking to overcome trickery that does not exist.

g. Avoid negative test items. These are too easily misread ("yes, I have no bananas"). Examine students on what to do, NOT

what NOT to do. Instruction reinforced in a positive manner is easier to apply. Negative test items are proper only when the student must know what to do and what NOT to do in order to avoid an undesirable consequence. For example, if the student has a buddy with a head injury and wants to ease his suffering, what should he avoid? He should NOT give the injured man morphine because the use of this drug could kill the injured person.

h. Avoid question tie-ins and cross keys. Each test item should stand alone. The answer to one test item should not give away the answer to another. Likewise, the answer to a test item should not depend on the ability to answer a previous test item correctly. Questions may refer to the same scenario or situation as long as the answers remain independent of each other.

i. Arrange test items in a logical sequence.

(1) Usually, the sequence of test items follows the sequence of subject matter presented during the instruction. This makes it easier for the student to keep his/her thoughts organized during the test rather than forcing him/her to shift back and forth from subject to subject.

(2) Tests may also be arranged according to the type of item. All multiple choice, all short answer, etc., will be grouped together. This grouping simplifies directions given to students and students are usually able to complete the test more quickly than they might if confronted with different types of test items.

(3) Test items may also be grouped according to the level of difficulty, that is, the easier ones first, the more difficult next, and the most difficult last. With this kind of test, students usually use their time more efficiently and have more time to answer the more difficult test items. Training developers should determine which is the best way to arrange test items.

j. Determine the length of the test. The training developer must choose the number and type of test items that will cover the instructional material adequately and treat the most essential and significant portions of instruction without weakening test reliability and validity. Other major factors governing the length of the test are the time that is available and learner fatigue. Written tests can be exhausting and time periods of more than two hours should be discouraged. As a guide, allow students about twice the time it takes a trainer to read through and complete the test.

k. Use as much of a map, schematic, illustration, photograph, or other visual aid as necessary for understanding the test item.

l. Test only one principle per item.

m. Ensure that only one response is correct.

n. Follow rules of grammar.

o. Write directions that include the time allowed, the manner in which students are to select and record answers, and the scoring system to be used. Practice test items may be provided to verify student understanding of the directions.

p. Avoid using absolutes such as "always" and "never."

q. Design a test to determine student learning, not to cause a student to fail.

r. Show the test items to several other people to rule out obvious faults which you have not perceived.

CHAPTER 3

STEPS IN WRITING A TEST

3-1. Assess the objectives. Each objective in a lesson is subject to being tested. If an objective is poorly written, it will be impossible to write a test that will properly test student achievement of the objective. The objective should be analyzed to determine the appropriate learning domain (cognitive, psychomotor, or affective). The conditions and standards should be reviewed to determine if they are realistic.

3-2. Assess the behavior required. Compare the behavior in the objective with the behavior required in the test item to determine if the behavior in the test item matches the behavior required in the objective. (i.e., if the verb used in the objective is "describe," the objective requires an essay test item; if the behavior is demonstrate, the objective requires a performance test.)

3-3. Construct a Test Item Pool. Write several items that could measure a single objective. The behavior and content of the objective will determine the type(s) and number of possible test items.

3-4. Select Test Items for a Given Test. Ensure that each block of instruction is sampled in the test.

3-5. Validate the Test(s). Validation usually consists of a minimum of three trials.

a. Try the test on a selected sample of the target population. Inform them that they are validating a test. Record any comments they may have for later consideration.

b. Administer the test to an entire class of the target population who are informed of the reason for their participation. Record comments.

c. Administer to an entire class without telling them they are validating the test. Record comments.

NOTE: Corrections should be made after each validation trial.

CHAPTER 4

GUIDELINES FOR WRITING TEST DIRECTIONS

4-1. Writing proper instructions is an essential part of any test. Instructions must be as clear, unambiguous, and as brief as possible. This applies to the general instructions given before the test and to the specific instructions immediately preceding the items to which they apply.

4-2. General Instructions. General instructions should include the following as a minimum:

- a. The objective of the test to include:

The purpose of the test.

(c) Time limits for the test. State if there is no time limit.

Description of test conditions.

Test standards.

- b. Special instructions such as:

"Do not talk to anyone during this test; raise your hand if you have a question."

or

"When you have completed the work at one station, proceed to the next station."

or

"When you have completed all requirements for this test, turn in your paper to the testing officer and remain in the waiting area until called by the instructor or testing officer."

4-3. Write understandable directions for each type of test item.

- a. Explain where answers are to be written/indicated:

Directly on the test paper.

In the column.

(3) On a separate sheet of paper (and whether student or instructor supplies the paper).

In the space provided, etc.

b. Explain how answers are to be written/indicated:

Circle or underline.

Rounded up/down or exact.

Length and content of answer.

4-4. Give an example of each type of question and the answer at the beginning of each section.

4-5. Specific Instructions. Sometimes specific instructions for each item are necessary. If the item requires special equipment or facility setups, special conditions, or specific standards which the test administrator must implement as a part of administering that item, specific instructions are required. If the item requires that special instructions be presented to the trainee in order for him/her to attempt it, specific instructions are required. When developing specific instructions, keep in mind the following points:

a. Specific instructions should be placed with the items to which they apply. Those parts of the specific instructions which the trainee should read are written into the item. Those parts which tell the administrator what to do should be included only in a separate "administrator's test copy."

b. Specific instructions should tell the trainee whether speed or accuracy is more important. Any time limits should be specified.

CHAPTER 5

GUIDELINES FOR WRITING MULTIPLE CHOICE TEST ITEMS

NOTE: Each chapter which follows deals with the development of specific types of test items and includes examples of each type of test item. There may be other ways of preparing items not included. The punctuation to be used with each type of test item will be discussed briefly, as applicable, after the example. Test items must be based on lesson objectives; therefore, the objective is provided first, then the test item is stated so that the relationship between objective and test item is clear. The objective would not normally be provided on the actual test. Sample directions for completing the test items are also included.

5-1. The decision to write a multiple choice test item is based on the behavior required by the objective.

a. Behaviors such as "select," "choose," and "match" are used in objectives which may use multiple choice test items.

b. The student is expected to select the answer to a question when given a list of possible answers.

c. This type of test item requires cognitive recall and recognition skills.

d. Students should be tested only on relevant information.

5-2. Stem -- contains the central problem and should follow these guidelines:

a. Wording should be clear and direct, focus on the action portion of the objective being tested, and be free of unnecessary information.

b. Terminology should be appropriate for the level of students being taught.

c. A problem should be stated so clearly that the student knows what is being asked before reading the alternatives.

d. All of the information needed to respond to the test item must be included. The stem should be written so that it can stand alone but should not be easily answered by the untrained student.

e. Sentences must be concise and grammatically correct.

f. Words in the stem should not be found in the answer or give clues to the answer; all repetitive phrases should be included in the stem.

g. Negative words should be avoided; however, if a negative is used it should be underlined for emphasis.

5-3. Responses/Distractors/Alternatives -- contain the correct and incorrect choices and should follow these guidelines:

a. Be complete.

b. Not be opposites, synonyms, or paraphrases of one another.

c. Not be answerable on the basis of common sense.

d. Avoid "all of the above," "none of the above," "a, b, and d above."

e. Be uniform in length and parallel in grammatical structure.

f. Include four logical responses which are common misconceptions of the logical solution to the problem. The correct answer MUST be the best answer.

g. Agree grammatically with the stem

h. Be specific, clear, and brief as possible

i. Be arranged in ascending or descending order with respect to numbers, length of wording, or chronology.

5-4. The location of correct responses should be varied; patterns should be avoided.

5-5. Advantages:

a. Easy to score by hand or machine.

b. Adaptable; can measure a variety of learning levels from simple to complex.

c. Familiar to students.

d. Can be completed in minimum time; it takes students less time to answer more questions; more information can be covered in a shorter period of time.

e. Students cannot bluff their way through the test as they may try to do in an essay test situation.

5-6. Disadvantages:

- a. May only measure recognition.
- b. Requires skill to develop test item.
- c. Difficult to develop legitimate distractors.
- d. Students are often able to guess the correct answer.
- e. May develop memorization instead of application skills

5-7. Multiple Choice Test Items: The multiple choice variety consists of an item stem followed by several responses. One of the responses is absolutely correct while the others are incorrect.

a. Objective: Given an NBC term and a list of definitions, select the definition of the term as it applies to the NBC environment IAW STANAG 2352.

b. Directions: Select your answer from the list of definitions which follows each question. Place the letter of the definition you have selected in the space provided on your answer sheet.

c. Test Item Example:

1. IAW STANAG 2352, what is the definition of the term "disinfection" as it applies to the NBC environment?
 - a. The act of depositing biological or chemical agents or radiological material on structures, areas, personnel, or objects.
 - b. The process of making any person/object/area safe by absorbing, destroying, neutralizing, or removing chemical or biological agents or by removing radioactive material.

- c. The act of depositing biological or chemical agents or radiological material on structures, areas, personnel, or objects.
- d. The act of inactivating the larger portion of microorganisms in or on a substance with the probability that all pathogenic bacteria are killed by the agent used.

5-8. The Incomplete Statement Variety of Multiple Choice: This type of test item consists of a stem containing a portion of the statement to be completed by the correct answer which the student must select.

a. Objective: Given an NBC term and a list of definitions, select the definition of the term as it applies to the NBC environment IAW STANAG 2352.

b. Directions: Select the answer which best completes the statement in each of the following items. Place the letter of the answer you have selected in the space provided on your answer sheet.

c. Test Item Example:

1. IAW STANAG 2352, the definition of the term "disinfection" as applied to the NBC environment is
 - a. the act of depositing biological or chemical agents or radiological material on structures, areas, personnel, or objects
 - b. the process of making any person/object/area safe by absorbing, destroying, neutralizing, or removing chemical or biological agents or by removing radioactive material.
 - c. the act of depositing biological or chemical agents or radiological material on structures, areas, personnel, or objects.
 - d. the act of inactivating the larger portion of microorganisms in or on a substance with the probability that all pathogenic bacteria are killed by the agent used.

OR

d. Objective: Given a list of various types of foods contaminated with a nuclear, biological, or chemical agent, and a list of methods used to treat contaminated food, match the contaminated food with the method used to decontaminate it IAW STANAG 2352.

e. Directions: Select the answer which best completes the statement in each of the following items. Place the letter of the answer you have selected in the space provided on your answer sheet.

f. Test Item Example:

1. In accordance with STANAG 2352, meat which has been contaminated with a nuclear agent is best decontaminated by
 - a. washing with soap and water.
 - b. washing with disinfectant.
 - c. boiling or soaking in water.
 - d. weathering.

g. Punctuation. The statement will be complete when the answer is selected. No punctuation is required at the end of the stem; a period should be used after each alternative.

CHAPTER 6

GUIDELINES FOR WRITING ALTERNATIVE CHOICE TEST ITEMS

6-1. An alternative choice (true-false, yes-no, agree-not agree) item consists of a declaratory statement or a situation that has only two choices. It is a two-response multiple choice test.

6-2. This type of question suffers from low reliability since the format encourages guessing and each choice has a fifty percent chance of being correct. It is often very difficult to construct questions that are absolutely true or absolutely false

6-3. Alternate choice test items may be used for quizzes but usually should be avoided on major tests and examinations.

6-4. Guidelines:

- a. Prepare test items in pairs.
- b. Write test items that test the objective and represent knowledge that is relevant. The test item should be written to preclude anyone with just common sense, but with no training, from answering it correctly.
- c. Avoid making the test item too obvious or too simple.
- d. Make test items challenging but not "tricky."
- e. Write statements that are entirely true or entirely false. There should be only one idea, fact, concept, etc., that determines the truth or falsity of the test item.
- f. Avoid making the test item statement too long. Long statements tend to obscure the basic true-false decision.
- g. Write so that the critical part of the test item is at the end of the statement.
- h. Avoid including words which are specific determiners/ absolutes: "All," "every," "never," "generally," "only," or "usually," "most," "many." These indicate a false response and an answer frequently hinges on these qualifiers.
- i. State questions positively. Avoid double negatives or statements where the correct answer combines awkwardly with the main statement (e.g., "yes, we have no bananas").

j. Design the test items to require the student to make application of the things learned.

k. Ensure that your teaching division allows this type of test item to be used before writing the test item.

l. Write a large enough number of items to ensure a good test. A good rule to follow is that the number of test items should be 2 to 4 times the number of minutes allowed for the test. Thus, a 10 minute quiz should have approximately 30 true-false items.

m. Vary true-false responses so that a pattern is not established.

n. Avoid:

(1) Developing questions by extracting statements verbatim from the text.

(2) Including two or more ideas in one test item.

6-5. Advantages:

- a. Easy to score. Scoring is very objective
- b. Used with a wide range of content.
- c. Can cover a large number of items in a short period of time.
- d. Is familiar to students.

6-6. Disadvantages:

- a. Easily leads to construction of trivial questions/statements.
- b. May be misleading.
- c. Difficult to construct completely true or false statements.
- d. Students have at least a 50 percent chance of guessing the correct answer.
- e. Reliability of results is questionable.

f. Test requires a large number of test items to effectively measure student learning.

g. May require memorization rather than application of student learning.

6-7. Alternative Choice (True-False) Examples: True-false variety. This type of test item consists of a declarative statement that is true or false.

a. Objective: Given an NBC term and a list of definitions, select the definition of the term as it applies to the NBC environment IAW STANAG 2352.

b. Directions: In the space provided on your answer sheet, state whether each of the definitions which follow the term being defined is true or false. Write the word "true" or "false" to indicate your answer. "T" or "F," etc., will be marked incorrect.

c. Test Item Example:

1. In accordance with STANAG 2352, the definition of the term "disinfection" as applied to the NBC environment is
 - a. the act of depositing biological or chemical agents or radiological material on structures, areas, personnel, or objects.
 - b. the process of making any person/object/area safe by absorbing, destroying, neutralizing, or removing chemical or biological agents or by removing radioactive material.
 - c. the act of depositing biological or chemical agents or radiological material on structures, areas, personnel, or objects.
 - d. the act of inactivating the larger portion of microorganisms in or on a substance with the probability that all pathogenic bacteria are killed by the agent used.

OR

d. Objective: Given a list of chemical agents and a list of their characteristic symptoms, match the symptom with the chemical agent which causes the symptoms IAW STANAG 2352.

e. Directions: In the space provided on your answer sheet, state whether each of the definitions which follow the term being defined is true or false. Write the word "true" or "false" to indicate your answer. A single letter, "T" or "F," etc., will be marked incorrect.

f. Test Item Example:

1. In accordance with STANAG 2352, early choking and stinging of the eyes and throat are symptoms of nerve agent contamination.

6-8. The Right-Wrong Variety of the Alternate Choice: This type of test item consists of a sentence, equation, or other expression that is to be marked right or wrong depending on whether it is correctly or incorrectly written.

a. Objective: Given an NBC term and a list of definitions, select the definition of the term as it applies to the NBC environment IAW STANAG 2352.

b. Directions: In the space provided on your answer sheet, state whether each of the definitions which follow the term to be defined is correct or incorrect. Write the word "correct" or "incorrect" (or "YES" or "NO") to indicate your answer. The single letter "C" or "I," etc., will be marked incorrect.

c. Test Item Example:

1. In accordance with STANAG 2352, the definition of the term "disinfection" as applied to the NBC environment is
 - a. the act of depositing biological or chemical agents or radiological material on structures, areas, personnel, or objects.
 - b. the process of making any person/object/area safe by absorbing, destroying, neutralizing, or removing chemical or biological agents or by removing radioactive material.

- c. the act of depositing biological or chemical agents or radiological material on structures, areas, personnel, or objects.
- d. the act of inactivating the larger portion of microorganisms in or on a substance with the probability that all pathogenic bacteria are killed by the agent used.

OR

d. Objective: Given a list of chemical agents and a list of their characteristic symptoms, match the symptom with the chemical agent which causes the symptoms IAW STANAG 2352.

e. Directions: In the space provided on your answer sheet, state whether the following definition is correct or incorrect. Write the word "correct" or "incorrect" (or "YES" or "NO") to indicate your answer. The single letter "C" or "I," etc., will be marked incorrect.

f. Test Item Example:

- 1. In accordance with STANAG 2352, early choking and stinging of the eyes and throat are symptoms of nerve agent contamination.

6-9. The Cluster Variety of the Alternate Choice: This variety consists of an incomplete stem with suggested completions. Each question is to be judged true or false.

a. Objective: Given a list of chemical agents and a list of dissemination mediums, match the chemical agent to its dissemination medium IAW STANAG 2352.

b. Directions: In the space provided on your answer sheet, state whether each of the following statements is true or false. Write the word "true" or "false" to indicate your answer. The single letter "T" or "F," etc., will be marked incorrect.

c. Test Item Example:

- 1. In accordance with STANAG 2352, nerve agents are disseminated through

- a. Vapors
- b. Fine powders.
- c. Liquid droplets.
- d. Aerosols

d. Punctuation. The statement will be complete when the answer is selected. No punctuation is required at the end of the stem and a period should be at the end of each alternative.

CHAPTER 7

GUIDELINES FOR WRITING MATCHING TEST ITEMS

7-1. A matching test consists of two sets of items with direct relationships.

7-2. Guidelines:

a. Place the stimulus (the "givens") items in the left column and the response (alternative) items in the right column.

b. Number the premises in the left-hand column.

c. Letter the responses in the right-hand column.

d. Develop several extra distractors in the response column to minimize the process of elimination. Items should be homogeneous (i.e., of a similar nature).

e. Directions must be clearly worded to specify the number of times the responses may be used (i.e., more than once, only once, etc.) and to clearly reflect how to match the numbers and letters.

f. Ensure the wording is clear in all premises and responses.

g. Ensure that the test item equates to the course objective it is testing.

h. Arrange the items in a logical order, alphabetical or numerical.

i. Limit the number of items. Fifteen is usually sufficient.

7-3. Advantages:

a. A good tool for testing association skills.

b. Relatively easy to construct; easy to score.

c. Not as easy to guess the correct answer.

d. Good for testing students' ability to match relationships between items.

7-4. Disadvantages:

- a. Usually limited to highly factual material which encourages memorization rather than application of information.
- b. Can encourage guessing.
- c. Requires skill to construct

7-5. Matching Test Item: The matching test item consists of two sets of items with a direct relationship. The student must recognize the relationship between the two items.

a. Objective: Given an NBC term and a list of definitions, match the definition to the NBC environment term to which it applies IAW STANAG 2352.

b. Directions: Match the items in Column 2 (lettered) with the items in Column 1 (numbered). Place the letter of the answer you have selected in the space provided on your answer sheet. Each of the items in Column 2 may be used more than once or not at all.

c. Test Item Example:

COLUMN 1 <u>Terms</u>	COLUMN 2 <u>Definitions</u>
1. Disinfection	a. A living animal or plant
2. Decontamination	b. A microorganism that causes disease
3. Endemic	c. Process of inactivating the larger portion of microorganisms
4. Biological Agent	d
5. Potable	
6.	

d. Test Item Example: True/False Variety of Matching

e. Directions: State whether the following statements are "True" or "False" in the space provided on your answer sheet. A single letter "T" or "F" will be marked incorrect.

1. IAW STANAG 2352, the definition of the term "disinfection" as it applies to the NBC environment is the act of depositing biological or chemical agents or radiological material on structures, areas, personnel, or objects.
- 2.

CHAPTER 8

GUIDELINES FOR WRITING SHORT ANSWER/COMPLETION/LIST TEST ITEMS

8-1. The short answer test item usually requires an answer that is longer than one word but not longer than one or two sentences. The short answer test item normally leaves a space for the answer to be inserted rather than using a blank which is used in the completion item.

8-2. Completion test items usually require only one word for an answer. The student responds to a direct question or incomplete declarative statement by writing an answer in the space (blank) provided.

8-3. Additional construction principles for the short answer and completion test items are basically the same.

8-4. Short answer and completion test items require that the student answers from memory (recall rather than recognition).

8-5. Guidelines:

a. Make a statement in which a critical key word or phrase is missing. Avoid omitting trivial words in the statement.

b. Avoid omitting verbs from the statement. Verbs should be included since it is easy to supply sensible alternate answers when a verb is required.

c. Word the statement so that the blank is near the end of the statement; avoid placing blanks near the beginning of the sentence.

d. Avoid copying statements verbatim from texts.

e. Allow for synonyms instead of requiring one "magic word" as the answer unless there is a good reason for requiring that one particular word. In this case, ensure that no other word could be a sensible answer.

f. Ensure that the test items test the objectives. Avoid counting spelling as a factor in scoring unless this was stated as a standard in the objective. If the objective requires correct spelling of the answer, make a provision to score the spelling requirement separately from the completeness requirement.

g. Avoid "Swiss cheese" (i.e., multiple blanks) test items where almost any combination of fill-in words could make the sentence a sensible one.

h. Make sure that inclusion of the missing word in the statement makes sense.

i. Avoid giving irrelevant clues ("a" or "an" or the exact number of spaces required to complete the answer) which give the answer away or make it easier for the student to guess an answer.

8-6. Advantages:

- a. Discourages guessing.
- b. Discriminates nonperformer from performer.
- c. Requires recall or production of answer.
- d. More objectives can be tested; amount and type of subject matter are only limited by test length.
- e. Test is taken rapidly.

8-7. Disadvantages:

- a. Tends to require pure memorization of facts.
- b. Requires skill in construction of test items.
- c. Requires decisions about partial or unexpected answers.
- d. Hand-writing may interfere with objective scoring.
- e. Difficult to evaluate application of learning.
- f. Spelling may interfere with objective scoring.

8-8. Guide for Writing Listing Test Items

a. Student will be expected to recall (memorize) the designated list. The training developer should be certain that the list is important enough to require memorization.

b. Be sure that the students are not asked to list trivia.

c. Ensure that the list reflects the objective it is testing.

d. Allow for reasonable synonyms or alternative responses unless there is a valid reason for requiring a specific word(s)

e. If the objective requires sequencing of the items listed, make a provision to score the sequencing requirement separately from the completeness requirement.

f. If there is a specific number of items to be listed, include this number in the directions.

8-9. Completion/Short Answer:

a. Objective: List the nine types of chemical agent detection equipment currently in use IAW STANAG 2352.

b. Directions: Answer the following questions in the space provided on your answer sheet. Each correct response will be awarded one point. Partial credit will NOT be awarded.

c. Test Item Example: The question variety:

1. What are the nine types of chemical agent detection equipment currently in use IAW STANAG 2352?

OR

2. What are six of the nine types of chemical agent detection equipment currently in use IAW STANAG 2352?

d. Objective: List the nine types of chemical agent detection equipment currently in use IAW STANAG 2352.

e. Test Item Example: The completion variety:

List the nine types of chemical agent detection equipment currently in use IAW STANAG 2352.

CHAPTER 9

GUIDELINES FOR WRITING ESSAY TEST ITEMS

9-1. Essays present a challenge to the instructor/test scorer. An essay test item requires the students to organize and write their thoughts about the subject. It is used to measure complex learning outcomes. Since people tend to think in different patterns and use different words to express themselves, no two answers will be the same.

9-2. An essay test usually requires answers of one or more paragraphs. Students will be expected to organize material, decide what is important, give descriptions, and/or discuss without being given alternate answers or prompts of any sort

9-3. Guidelines:

a. Construct simple and clear test items based on the lesson objectives so that the student is directed to the desired response.

b. State the problem as clearly as possible. Define the direction and scope of the desired answer. Give the students as much information about the requirements of a correct response as possible. Enumerate what points are to be included in the response, if possible.

c. Avoid general, all-encompassing test items.

d. Indicate the amount of time to be spent on each item.

e. Include a word/space limit (delimiter).

f. Prepare the answer or expected response as the test item is constructed. Break each question down into its different parts and identify the major points which must be covered by the students' answers. Specify exactly how much credit will be given for each part of the answer. The predetermined system for evaluating the student should be usable by several scorers and produce the same results.

g. Avoid being influenced by writing style unless it's included in the behavior.

h. Make provisions for anonymous grading.

i. Present scenarios/situations which require application of knowledge and/or experience.

j. Before scoring the tests, select a few answers at random to determine how they compare to your ideal answer. If you see that students are answering a different question than the one you intended, change the criteria for a correct answer. Once you begin grading do not change the criteria again.

k. Score all answers to one question at the same time and in one sitting before moving on to another question.

l. Provide constructive feedback to ensure students understand and learn from their mistakes.

m. The following are examples of behaviors which are used to develop essay objectives/test items:

(1) Compare: Show the similarities between two events, periods, people, and ideas.

(2) Contrast: Show the differences between two or more events, periods, people, and ideas.

(3) Describe: Give an account of or present the characteristics of.

(4) Discuss: Consider or examine by argument, comment, etc. Give the pros and cons of, write about, debate, etc., of an issue, event, or person.

(5) Evaluate: Make a judgment, give an opinion, supply supporting reasons.

(6) Explain or interpret: Support, clarify, or qualify a given generalization with specific facts and ideas.

(7) Prove: Confirm, substantiate, verify or demonstrate the authenticity or truth of a statement, or the reasons for a historical event.

9-4. Advantages:

a. Assesses the student's ability to organize and compose or construct answers.

b. Permits the student to develop or create a product or answer from hypothetical or similar situations.

c. Serves as a great source of feedback to the instructor and the instruction.

d. Can be used to measure many aspects of student learning such as a student's writing ability, verbal self-expression, problem-solving skills, etc.

e. Students cannot "guess" the right answer.

Disadvantages:

a. Difficult to develop complete answers for each test item.

b. Extremely difficult to score reliably. It is the least objective of all types of testing methods.

c. The test writer must determine exactly what information is being presented and must provide a comprehensive answer with correct information.

d. It requires the grader/instructor to determine the value of additional information presented by the students.

e. It is time-consuming to complete and to score.

f. Poor hand-writing or writing style may interfere with objectivity of scoring.

g. It may favor verbally adept student

h. It is not always easy to detect bluffing.

NOTE: Multiple choice and/or matching test items may be used as an alternate method of testing essay test items. Caution must be taken to ensure that the information tested accurately reflects the content and intent of the objective.

Essay Test Items:

a. Objective: Describe the hazards associated with nuclear fallout IAW STANAG 2352.

b. Directions: Write the answer to the following essay questions on your answer sheets. Limit your answers to the number of words cited in each question.

c. Test Item Example:

In less than 100 words, describe the hazards associated with nuclear fallout IAW STANAG 2352.

OR

d. Objective: Discuss the defensive countermeasures for the protection of food and water from NBC contamination to include detection and purification of contaminated water and the reclamation of contaminated food, forage, and grain IAW STANAG 2352.

e. Test Item:

Use no more than 500 words to discuss the defensive countermeasures for the protection of food and water from NBC contamination to include the detection and purification of contaminated water and the reclamation of contaminated food, forage, and grain IAW STANAG 2352.

CHAPTER 10

GUIDELINES FOR WRITING PERFORMANCE TEST ITEMS

10-1. Task-oriented, hands-on training: Performance tests are formal assessments of the student's ability to perform psychomotor tasks in an environment which closely resembles that of the real world. The test is based on observation of his/her performance (process, product, or combination) evaluated against a predetermined standard.

10-2. Performance tests consist of the measurement of physical or motor skills but contain elements of all three learning domains.

10-3. Performance tests let the instructor know that students can actually perform their hands-on tasks before they go out into the field to work on real patients/equipment. (The fact that students can answer questions on a knowledge-based test does not always indicate that they can actually perform the required psycho-motor steps and maintain the proper attitude.)

10-4. Types of performance tests:

a. Process: Measures sequence or clusters of steps/actions that students perform to accomplish task.

b. Product: Used to measure a finished product - a tangible item - resulting from a process.

c. Combination: Used to measure the process and the product.

10-5. Guidelines:

a. Ensure that the required performance produces some definable end product or reveals some specific observable process which is based on the lesson objective.

b. Base the required performance upon clear, easily applied standards which are easy for almost anyone to evaluate accurately.

c. Provide an evaluation checklist which specifies all required elements for the performance in the required sequence, and which permits separate grading of each element.

d. Develop a clear and easy to use grading system. Ensure that the grading system provides for an accurate pass/fail determination and specific feedback to the student as to his/her strong and weak points. Checklists are often used to evaluate hands-on student performance.

e. Ensure that the test item does not require equipment or personnel support beyond that which can be provided on a routine basis.

f. Ensure that the test item parallels job conditions as much as possible.

g. Ensure that the performance can be completed in the allotted test time by most students. If time standards are relevant, ensure that they are included in the instructions to the students and in the grading system.

h. Ensure that the instructions are clear and state exactly what the student is expected to do.

i. Ensure that the directions for test administrators permit the test item to be administered consistently and specifies all necessary procedures and items of student equipment.

j. Provide instructions which inform the student of the grading standards which will be used in evaluating his performance.

k. Show this test item to several other people to rule out obvious faults which you did not perceive.

10-6. Advantages:

a. Confidence in a student's ability is increased when the student demonstrates a skill or behavior in a realistic setting.

b. May demonstrate affective behaviors which are difficult to assess otherwise.

c. Allows role playing.

d. Allows first-hand observation of student performance.

10-7. Disadvantages:

a. Resource intensive

- b. May be impossible to test in a realistic setting.
- c. Subjectivity may interfere with evaluation
- d. Student anxiety may adversely affect performance.

10-8. Performance Test Item Sample

a. Directions: Include all parameters for completion of the task. Ensure that the student is provided with all necessary equipment ("givens"), etc., and that the student knows the standards required for successful completion (GO) of the task (i.e., 70% of the steps must be successfully completed). If any steps of the tasks are critical, these steps should be identified by an asterisk throughout training and testing. The critical steps of the task must be successfully completed to earn a "GO" for the task regardless of the percentage of noncritical steps which are completed.

b. Objective: Given a simulated nerve agent casualty and three sets of injectors, administer the three nerve agent antidote injections within 2 minutes IAW _____.

c. Test Item Example:

Use the three sets of injectors to administer three nerve agent antidote injections within 2 minutes and IAW _

GO NO GO

1. Remove one set of antidote auto-injectors from inside the pocket of the casualty's mask carrier.
2. Grasp the injector set by the plastic clip with your dominant hand.
3. Hold the injector set in front of your body at eye level.
4. Hold the large injector on top.
5. Grasp the small auto-injector with the thumb and first two fingers of your opposite hand.

GO NO`GO

6. Pull the injector out with a smooth motion to arm the injector.
(Green end should not be pressed until the injection is ready to be administered.)
7. Form a fist around the auto injector.
8. Place the green (needle) end of the injector against the casualty's outer thigh muscle.
(If the casualty has a thin lateral thigh muscle, administer the injection(s) in the outer quadrant of the buttocks.)
9. Avoid items in the casualty's trouser pockets when placing the needle against the casualty's outer thigh muscle.

Apply a firm, even pressure to the injector until it functions (clicks).

11. Hold it in position for 10 seconds.

Remove the atropine auto-injector and place the expended auto-injector between the little finger and the ring finger of the holding hand.

Administer the Pralidoxime Chloride (2 PAM Cl), etc...

Continue to list the steps which the student must complete to accomplish the task.

CHAPTER 11

GUIDELINES FOR FORMATTING A TEST

- 11-1. Be consistent in presenting an item type (i.e., if using multiple choice, include 4 responses consistently throughout; avoid having 3 responses in one item and five in another, etc.)
- 11-2. Keep all parts of a test on the same page. Do not separate a question or group of questions from a reading passage, table, graph, or drawing that is required to complete the test item.
- 11-3. Avoid crowding too many test items on one page of the test.
- 11-4. Group test items by type. Put all completion test items together, all multiple choices together, etc.
- 11-5. Begin with easier test items to give students the opportunity to "break into" the test.
- 11-6. Arrange the test items within each of the question types in ascending difficulty, proceeding from the simple to the more complex.
- 11-7. Revise all test items that provide answers to other test items.
- 11-8. Number all test items in consecutive order.
- 11-9. Single-space each situation, stem, and answer choice but double space between each situation, stem, and answer choice.
- 11-10. Place figures, illustrations, and situations before the group of test items to which they pertain.
- 11-11. Ensure test items using figures, illustrations, and situations clearly indicate how to use the material in answering the test item.
- 11-12. Try to keep illustrations no larger than 6 X 8 inches.

CHAPTER 12

GUIDELINES FOR SCORING TESTS

12-1. Use the following to score individual items for all tests.

- a. Correct = No marking(s) made.
- b. Incorrect = X.
- c. Omission = O.

12-2. Analyze test results. Consider the items below.

- a. Number of GO/NO GOs.
- b. Item analysis. For each item determine
 - the number of student selections for each response.

(2) the difficulty (percentage of students answering correctly).

the effectiveness of distractors.

12-3. Reasons for missing answer. For each item consider if the problem is

- a. student error.
- b. printed material incorrect.
- c. presentation of material incorrect.

12-4. Scoring Essay Tests. Essay tests require some special considerations when scoring:

a. Prepare an outline of acceptable answers. More than one person might be scoring the test.

b. List other factors that might be taken into consideration such as the following:

Additional information.

Handwriting.

(3) Spelling.

Punctuation.

Style and organization.

Neatness.

12-5. Inform the class of the other scoring factors

12-6. If possible, student identity should be unknown.

12-7. Score all answers pertaining to one test item at one sitting.

12-8. Shuffle papers before grading next test item.

(HSHA-EE)

FOR THE COMMANDER:

OFFICIAL:


CLODETH C. FINDLAY
CPT, MS
Adjutant General

CHARLES E. DYER, II
Colonel, MS
Secretary of the General Staff

DISTRIBUTION:

E, plus:
200-HSMC-IAP
150-HSHA-M
50-HSMC-FT
20-HSHA-E
1-HSHA-ME