

Multiple Choice Question and Performance Examination Item Types

Question: *What are multiple choice questions and performance examination item types and how are they used in licensure and certification examinations?*

Answer:

Multiple Choice Question Item Types

Multiple choice questions are the most commonly used type of test items employed in licensure and certification examinations. Use of multiple choice questions have been shown to be valid and reliable as well as a time and cost-effective means of testing knowledge and its application. They permit the testing of a wide range of discrete knowledge and skill components.

The most common type of multiple choice question consists of two basic parts: a stem and two or more choices that include a single best answer.

The stem often consists of an introductory question, statement or case which is followed by a lead-in question. The stem is generally followed by choice consisting of up to five response options (one correct or best answer and other choices called distractors).

The development of effective multiple choice questions is a craft that requires knowledge of the principles of item writing and a great deal of practice and time. Most testing agencies using multiple choice questions provide item writers with detailed training on writing effective multiple choice questions. Important principles in multiple choice item development include, but are not limited to the following:

1. Each item should focus on one objective that is part of an overall test blueprint. What do you want to find out about the test-taker?
2. Items should focus on application of knowledge and problem solving, rather than on recall of knowledge or facts
3. The item should reflect good test practices, to avoid error in scores caused by confusion, ambiguity or factors unrelated to the objective being tested, e.g.:
 - i) appropriate language should be used
 - ii) the stem should pose only one clear, succinct, unambiguous question
 - iii) the test taker should be able to answer the question by reading the stem, without looking at the choices
 - iv) the stem should be positively stated
 - v) the choices should be grammatically consistent with the stem and of similar length, structure and concept
 - vi) the distractors (incorrect answers) should be concise, comparable, plausible and related to common errors, misconceptions or unsafe practices
4. Common flaws in item writing include negatively stated questions (e.g., "all are true except"), the use of implausible distractors, use of inappropriate language and vocabulary, grammatical errors and use of distractors which include the use of "all of the above", "none of the above", "never" and "always"

Innovative Multiple Choice Item Formats: Innovative types of multiple choice question formats include the use of multiple-multiple choice questions (i.e., with more than one correct answer (examples are Key Feature and Pick-N items), situational sets or scenario-based items, and extended matching questions (i.e., items that are organized into sets that use one list of options for all items in the set, thus allowing for the choice of multiple correct answers). All require the same careful development and review process as is used in single-answer multiple-choice items.

Limitations: Multiple choice questions may be less valuable to test some important behavioral professional skills such as communication and interpersonal skills. They may be less appropriate to evaluate skills which require integration of clinical knowledge with communication/interpersonal and problem-solving skills. As a result of this limitation, some licensure and certification agencies are utilizing performance-based examination items in addition to multiple choice examinations.

Performance Exam Item Types

A performance assessment consists of one or more tasks, a specification of the conditions governing the performance of the tasks, and a method of scoring the result. There are several examples of well-designed performance-based assessments that are being used today. All “attempt to emulate the context or conditions in which the intended knowledge or skills are actually applied” (Standards for Educational and Psychological Testing, 1999). However, much care is required in the design, administration, and scoring rubrics to ensure good measurement practices are followed. The format known as an “objective structured clinical examination” (OSCE) is an example of a recognized, valid and reliable method for assessment of integration of knowledge, skills, and abilities. In general, OSCEs consist of a series of stations (complex tasks or items) through which all candidates rotate on a timed basis. Each station simulates a practice-based situation that is commonly encountered and/or critical in nature. The candidate is required to perform specific roles and functions to complete the task or address the problem depicted in the simulation. A trained assessor is frequently present in the station to provide real-time, direct observation and assessment of each candidate’s performance. Performance checklists and/or global rating scales based on standardized criteria are used to record and assess the candidates’ performance. Non-interactive (or quiet) station items may be used to assess candidates’ written responses to tasks or problems that do not require verbal communication. The development of station items require considerable time and expertise.

As a performance-based tool, OSCEs have demonstrated advantages over other traditional forms of assessment, such as multiple-choice questions or oral examinations. Communication and interpersonal skills, ethical and professional judgement, as well as complex problem identification and resolution skills may be assessed more effectively and efficiently through a well-designed OSCE than through other testing. By virtue of its performance-based nature, the OSCE is well suited to complement traditional paper-based, computer-based, or other testing methods. OSCEs are often used in health professions for licensure and certification purposes to assess how candidates perform the most frequent and critical situations that arise in practice and to assess their competence in these contexts.

Limitations: The cost and resources in developing and administering performance assessments limit their use. They are less efficient and effective in testing discrete knowledge components for feedback purposes in ensuring that candidates have the breadth of knowledge to practice safely.