Evaluations and Outcomes

Author:
Julia Thompson, PhD, RN, CNS, CNOR
Administrative Director, Research and Sponsored Programs
Harris County Hospital District
Houston, Texas

Reviewer:
Diana Beck, MSN, RN, CNOR
Perioperative Education Specialist
St. Mary’s Good Samaritan, Inc.
Centralia, Illinois
Adjunct Faculty
College of Nursing, University of Missouri, St. Louis
St. Louis, Missouri

CCI
Personal Commitment to Extraordinary Care®
Unit 1:
Introduction

This module is designed for nurses who currently function in an educator role, those identified as emerging educators, and those who anticipate filling educator roles in the future and are interested in developing knowledge of evaluation in nursing education. Nurse educators are responsible for appraising learning of staff, the success of educational program outcomes, and evaluation of their own personal teaching abilities. The content in this module provides an overview of the evaluation process, measurement, testing, assessment, and other concepts by which nurse educators can evaluate outcomes.

Evaluation is an important part of nursing education and nursing practice as it provides a means for ensuring accountability for the quality of education and services provided. Nurses evaluate interventions to determine whether they are helping achieve patient outcomes. In nursing education, educators assess learner progress to determine if they are achieving the objectives from an in-service or orientation program. All nursing professionals judge performance and determine competency to practice. Evaluation should be a part of all educational activities, and all educators should be knowledgeable about the process.

Objectives

Upon completion of this module, the participant should be able to:

1) Define evaluation.

2) Discuss the process of evaluation.

3) Select the appropriate evaluation tool based on the identified audience and outcomes.

4) Use guidelines for writing test questions to develop an effective test.

ANCC Accreditation Statement

This continuing nursing education activity has been approved by the Association of periOperative Registered Nurses (AORN), an accredited approver by the American Nurses Credentialing Center’s (ANCC) Commission on Accreditation. Activities that are approved by AORN are recognized as continuing education for registered nurses. This recognition does not imply that AORN or the ANCC Commission on Accreditation approves or endorses any product included in the presentation.
A popular method for evaluating learning outcomes is the use of written tests. Tests should produce information that is relatively objective and quantifiable. To construct effective tests, educators must have:

- knowledge of the content to be tested,
- knowledge of the level of the learner,
- the ability to communicate clearly in writing, and
- the willingness to spend the time necessary to produce a good test.

Tests may already have been developed for frequently taught topics and competencies, and the first strategy is to improve existing test items. The educator should always take the test to check for errors in content, wording, numbering, and unclear language. As, Bs, Cs, and Ds should be randomly allocated for correct answers. Stems ending with a question mark should have capitalized responses that end with a period; all stems that are open-ended or fill in the blank should have responses that begin with a lower case letter and end with a period (Aucoin, 2005). If staff members repeatedly answer the same question incorrectly, the educator should investigate possible reasons and modify the question as needed.

The following steps have been identified in constructing a written test (Alspach, 1995):

1. Based on the objectives for the activity, identify the instructional outcomes to be evaluated.
2. Determine the amount of time available for testing. A good rule of thumb is one question per minute.
3. Decide on the type(s) of test item(s) to be used (e.g., true/false, multiple choice, matching, fill in the blank, short story, essay).
4. Calculate the total number of test items that can be used in the time available for testing.
5. Calculate the number of test items that should be allocated to each instructional outcome.
6. Designate item numbers for the test.
**Item Writing**

Writing items for true-false, multiple-choice, and short answer tests requires time and skill. The definition, guidelines for writing, and examples of each of these types of test items follow.

**True-False Items**

A true-false question consists of a declarative statement followed by an opportunity for the learner to determine whether the statement is true or false. There are only two possible responses for each item, and the learner chooses between them.

True-false questions are well suited for testing learner recall or comprehension. Learners can generally respond to many questions, covering a lot of content, in a fairly short amount of time. From the educator’s perspective, these questions can be written quickly and are easy to score. Because they can be objectively scored, the scores are more reliable than for items that are at least partially dependent on the educator’s judgment.

*Guidelines for Writing True-False Items:*

- Statements should be relatively short and simple.
- Make sure the statement is clearly true or false.
- Avoid the use of absolute terms such as “all” or “always,” which make the items false. Likewise, qualifiers such as “sometimes” or “typically” would make the item true.
- Avoid negative and double negative statements.

*Examples:*

1. Straight Mayo scissors are used to cut delicate tissue. **T** **F**
2. Another name for a transverse surapubic incision is Pfannensteil. **T** **F**

**Multiple-Choice Items**

A multiple-choice question consists of a statement or question (the stem), followed by a number of possible responses or answers (the options or distractors). Multiple-choice is the most common type of test item.

*Guidelines for Writing Multiple-Choice Items:*

- Present a single problem in the stem. The answer should be obvious without looking at the options.
• Write the stem in a precise manner. If it is too complex, the learner will spend too much time trying to decipher it.

• Offer at least four options for each item.

• Be sure there is only one correct answer.

• Offer distractors that are definitely wrong.

• Be aware of using answers such as “none of these” or “all of the above.”

• All of the information needed to answer the question should be provided in the stem.

Examples:

1. The penetration of moisture through a bacterial barrier or vice versa is known as:
   a. transmission.
   b. fall out.
   c. strike-through.
   d. denaturation.

2. What is the most important step in the sterilization process?
   a. Air removal.
   b. Length of time.
   c. Moisture content.
   d. Decontamination.
   e. Pressure.

Short Answer

The short answer item requires the learner to produce an answer. Short answer items can be written in one of two forms — as a direct question or as a sentence having one or more important words deleted. The direct question requires the learner to supply a word, phrase, or a few sentences. The incomplete sentence usually requires the learner to supply only a word, a phrase, or a number that completes the statement.

Guidelines for Writing Short Answer Items:

• Write completion items that can be answered with one word, number, or phrase.

• Phrase the stem so that there can be only one unique word or phrase that is the correct response.

• Avoid giving grammatical clues to the answer (e.g., use of a or an).

• Keep the length of the blanks the same.
Unit 3: Tests

- Place blanks at the end of the sentences.

Examples:

1. List the advantage(s) of using atraumatic suture.

2. Sutures that are assimilated by body enzymes during the healing process are classified as _________.

Grading

Grading involves quantifying data and assigning value. Grades provide both feedback and motivation for learners. Passing or “cut” scores are typically set at a minimum of 80–85%, and may be predetermined by the facility or the organization providing the learning activity. Due to expert knowledge of the learner population, the content, and the consequences if the learner’s performance on the test translates to unsafe patient care practices, the educator may set the passing score. The educator may determine that for some items 100% accuracy is vitally important, such as with drug calculation problems, and put additional emphasis on those types of questions.

One method of assigning grades for clinical practice is the use of the “pass/needs additional work” or “met/not met” option. A passing score would consist of a minimum requirement for each behavior being evaluated. When clinical practitioners are expected to demonstrate achievement of objectives measured in a practice setting, pass or needs additional work is an appropriate grading method. Pre- and post-tests may be used to help educators determine baseline knowledge prior to an educational activity, and then evaluate the effectiveness with which staff members have met the objectives of the educational programs.

In the clinical setting, the most common method of determining the cut score is through informed judgment. Stakeholders (staff members, patients, nursing supervisors, human resources, legal representation) should all be considered when determining a passing score. The consequences of misclassification (e.g., if a person passes the test but is not competent, or does not pass the test but is competent) and whether unsuccessful test takers are allowed to re-take the test or offered opportunities for remediation, are additional issues that must be considered when developing an exam (Azzarello, 2003).

Reliability and Validity

As with using any measuring device, it is important to ensure validity and reliability of instruments used for evaluation. Whatever the purpose of the evaluation, the method used should satisfy the following three requirements:

1. Validity: Does it measure what it is supposed to measure?

2. Reliability: Does it produce consistent results?

3. Practicality: Is it practical in terms of time and resources?
For example, a test of blood sugar would be invalid if it actually measured blood alcohol. It would be unreliable if the same sample of blood gave different results on repeated analyses. It would be impractical if it took two laboratory technicians three hours to perform at a cost of $500 per test.

Validity

Validity refers to whether a test or evaluation instrument accurately measures what it is supposed to measure. The validity of any test is always judged in reference to a specific purpose. An evaluation tool designed to measure ability to recall facts should do just that. One designed to measure skills should measure the outcomes of the educational objectives.

Imagine a bathroom scale that consistently tells you that you weigh 130 pounds. The reliability (i.e., consistency) of this scale is very good, but is not very accurate (i.e., valid), because you actually weigh 145 pounds.

Reliability

Reliability is the consistency of measurement, or the degree to which an instrument measures the same way each time it is used under the same condition with the same subjects. A measurement is said to be reliable if it is reproducible. A measure is considered reliable if a person’s score on the same test given twice is similar. It is important to remember that reliability is not measured, it is estimated. Though reliability is initially less important than validity, it should remember that an unreliable assessment cannot be valid.

Another way to think of reliability is to imagine a kitchen scale. If you weigh five pounds of potatoes in the morning, and the scale is reliable, the same scale should register five pounds for the potatoes an hour later (unless, of course, you peeled and cooked them). Likewise, instruments such as a test should be reliable. It should not make any difference whether the learner takes the test in the morning or afternoon; on one day or the next.

Reliability and validity form the basis for all assessments. A test that is valid and reliable measures what it intends to measure and with a degree of accuracy that can be trusted. Considering the serious decisions that are made about learners based on tests, educators should have a working knowledge of the concepts of reliability and validity. Understanding these concepts provides educators with a framework both for developing more effective tests and for accurately interpreting and using test results.

Practicality

Also important is the practicality of evaluation procedures. Measurement procedures should be efficient and economical (Oermann & Gaberson, 2006). A test should be easy to construct, use, administer, and score; not take too much time away from other educator activities; and be cost-effective.

Bloom’s Taxonomy

Methods and procedures for evaluating learning should be well articulated and directly linked to the stated learner objective. Clearly stated behavioral objectives are vital to evaluating the outcome(s) of instruction.
References


Additional Resources

Web Sites

ERIC (Educational Resources Information Center)
http://www.eric.ed.gov/
National League for Nursing
http://www.nln.org/index.cfm

National Library of Medicine for Health Care Professionals

National Nursing Staff Development Organization
https://www.nnsdo.org/

Microsoft Office Online—provides tutorials for developing charts and graphs.
Click on Help and How To, to get started.

Journals

Journal of Nursing Education
International Journal of Nursing Education Scholarship
International Journal of Training and Development
Journal of Continuing Education in Nursing
Journal for Nurses in Staff Development
Nurse Educator

Books


