SPRINGER PUBLISHING COMPANY

SIXTH EDITION

Evaluation and Testing in NURSING EDUCATION

Marilyn H. Oermann
Kathleen B. Gaberson

EVALUATION AND TESTING IN NURSING EDUCATION

Marilyn H. Oermann, PhD, RN, ANEF, FAAN, is the Thelma M. Ingles Professor of Nursing at Duke University School of Nursing, Durham, North Carolina. She is an author or coauthor of 21 books and many articles on evaluation, teaching in nursing, and writing for publication as a nurse educator. She is the Editor-in-Chief of *Nurse Educator* and the *Journal of Nursing Care Quality* and past editor of the *Annual Review of Nursing Education*. Dr. Oermann lectures widely on teaching and evaluation in nursing.

Kathleen B. Gaberson, PhD, RN, CNOR, CNE, ANEF, is the owner of and principal nursing education consultant for OWK Consulting, Pittsburgh, Pennsylvania. She has over 35 years of teaching and administrative experience in graduate and undergraduate nursing programs. She is a coauthor of 10 nursing education books and an author or coauthor of numerous articles on nursing education and perioperative nursing topics. Dr. Gaberson presents and consults extensively on nursing curriculum revision, assessment and evaluation, and teaching methods. Dr. Gaberson is the Associate Editor for Research of the *AORN Journal* and also serves on the *Journal* Editorial Board.

EVALUATION AND TESTING IN NURSING EDUCATION

Sixth Edition

Marilyn H. Oermann, PhD, RN, ANEF, FAAN
Kathleen B. Gaberson, PhD, RN, CNOR, CNE, ANEF

SPRINGER PUBLISHING COMPANY

Springer Publishing Company, LLC
11 West 42nd Street
New York, NY 10036
www.springerpub.com
http://connect.springerpub.com/home

Acquisitions Editor: Adrianne Brigido
Compositor: diacriTech

ISBN: 978-0-8261-3574-2
ebook ISBN: 978-0-8261-3575-9
Instructor's Manual ISBN: 978-0-8261-3576-6
Instructor's PowerPoints ISBN: 978-0-8261-3577-3
DOI: 10.1891/9780826135759

Qualified instructors may request supplements by emailing textbook@springerpub.com

19 20 21 22 23 / 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data
Library of Congress Control Number: 2019917090

Printed in the United States of America.

CONTENTS

PART IV. Clinical Evaluation

PART V. Issues Related to Testing and Evaluation in Nursing Education

Appendices

PREFACE

All teachers at some time or another need to assess learning. The teacher may write test items; prepare tests and analyze their results; develop rating scales and clinical evaluation methods; and plan other strategies for assessing learning in the classroom, clinical practice, online courses, simulation, and other settings. Often teachers are not prepared to carry out these tasks as part of their instructional role. This sixth edition of *Evaluation and Testing in Nursing Education* is a resource for teachers in nursing education programs and healthcare agencies; a textbook for graduate students preparing for their role as a nurse educator; a guide for nurses in clinical practice who teach others and are responsible for evaluating their learning and performance; and a resource for other healthcare professionals involved in assessment, measurement, testing, and evaluation. Although the examples of test items and other types of assessment methods provided in this book are nursing oriented, they are easily adapted to assessment in other health fields.

The purposes of this book are to describe concepts of assessment, testing, and evaluation in nursing education and prepare teachers for carrying these out as part of their roles. The book presents qualities of effective assessment procedures (reliability, validity, and usability); how to plan for testing, assemble and administer tests, and score tests; how to write all types of test items and develop assessment methods; strategies for assessing higher level learning; and testing and evaluation in online courses and programs. The book describes the evaluation of written assignments in nursing, the development of rubrics, clinical evaluation, methods for evaluating clinical performance, and using simulation and objective structured clinical examinations (OSCEs) for evaluation. A new chapter prepares educators to analyze the performance of the test as a whole and of individual test items and to interpret test scores; this chapter includes many examples and exhibits to help readers understand these analyses and make informed decisions about their tests. This edition also examines the social, ethical, and legal issues associated with testing and evaluation in nursing; grading; and program evaluation (with a new section on accreditation of nursing education programs). The Appendices provide a quick reference guide for writing different types of test items (with examples) and

other testing resources. The content is useful for teachers in any setting who are involved in evaluating others, whether they are students, nurses, or other types of healthcare personnel.

Chapter 1 addresses the purposes of assessment, testing, measurement, and evaluation in nursing education. Differences between formative and summative evaluation and between norm-referenced and criterion-referenced measurements are explored. Because effective assessment requires a clear description of *what* and *how* to assess, the chapter describes the use of outcomes for developing test items, provides examples of outcomes at different taxonomic levels, and describes how test items would be developed at each of these levels.

In Chapter 2, qualities of effective assessment procedures are discussed. The concept of assessment validity, the role of reliability, and their effects on the interpretive quality of assessment results are described. Tests and other assessment instruments yield scores that teachers use to make inferences about how much learners know or what they can do. *Validity* is the adequacy and appropriateness of those interpretations about learners' knowledge or ability based on those scores. Current ways of thinking about reliability and its relationship to validity are explained. Also discussed in Chapter 2 are important practical considerations that might affect the choice or development of tests and other instruments.

Chapter 3 describes the steps involved in planning for test construction, enabling the teacher to make good decisions about what and when to test, test length, difficulty of test items, item formats, and scoring procedures. An important focus of the chapter is how to develop a test blueprint and then use it for writing test items; examples are provided to clarify this process for the reader. Broad principles important in developing test items, regardless of the specific type, are described in the chapter.

There are different ways of classifying test items: One way is to group them according to how they are scored—objectively or subjectively. Another way is to group them by the type of response required of the test-taker—selected or constructed response—which is how we organized the chapters. Selected-response items require the test-taker to select the correct or best answer from options provided by the teacher. These items include true–false, matching, multiple choice, and multiple response. Constructed-response items (fill-in-the-blank and essay) require the test-taker to supply an answer rather than choose from options already provided. Chapters 4 to 6 discuss these test items.

A true–false item consists of a statement that the student judges as true or false. In some forms, students also correct the response or supply a rationale as to why the statement is true or false. True–false items are most effective for recall of facts and specific information but may also be used to test the student's comprehension of the content. Chapter 4 describes how to construct true–false items and different variations, for example, correcting false statements or providing a rationale for the response, which allows the teacher to assess if the learner understands the content. Chapter 4 also explains how to develop matching exercises. These consist of two

parallel columns in which students match terms, phrases, sentences, or numbers from one column to the other. Principles for writing each type of item are presented, accompanied by sample items.

In Chapter 5, the focus is on writing multiple-choice and multiple-response items. Multiple-choice items, with one correct answer, are used widely in nursing and other fields. This format of test item includes an incomplete statement or question, followed by a list of options that complete the statement or answer the question. Multiple-response items are designed similarly, although more than one answer may be correct. There are three parts in a multiple-choice item, each with its own set of principles for development: (a) stem, (b) answer, and (c) distractors. In Chapter 5, we discuss how to write each of these parts and provide many examples. We also describe principles for writing multiple-response items, including the format used on the NCLEX® (National Council Licensure Examination).

Short-answer (fill-in-the-blank) items can be answered by a word, phrase, or number. One format presents a question that students answer in a few words or phrases. With the other format, completion or fill-in-the-blank, students are given an incomplete sentence that they complete by inserting a word or words in the blank space. On the NCLEX, candidates may be asked to perform a calculation and type in the number or to put a list of responses in proper order. In Chapter 6, we describe how to write different formats of short-answer items. We also explain how to develop and score essay items. Essay items provide an opportunity for students to select content to discuss, present ideas in their own words, and develop an original and creative response to a question. We provide an extensive discussion on scoring essay responses and on developing rubrics.

With higher level thinking, students apply concepts and other forms of knowledge to new situations, use that knowledge to solve patient and other types of problems, and arrive at rational and well thought-out decisions about actions to take. The main principle in assessing higher level learning is to develop test items and other assessment methods that require students to apply knowledge and skills in a *new* situation; the teacher can then assess whether the students are able to use what they have learned in a different context. Chapter 7 presents strategies for assessing higher levels of learning in nursing. Context-dependent item sets or interpretive exercises are discussed as one format of testing appropriate for assessing higher level cognitive skills. Suggestions for developing these are presented in the chapter, including examples of different items. Other methods for assessing cognitive skills in nursing also are presented in this chapter: cases, case studies, unfolding cases, discussions using higher level questioning, debates, video clips, and short written assignments.

Chapter 8 focuses on developing test items that prepare students for licensure and certification examinations. The chapter begins with an explanation of the NCLEX test plans and their implications for nurse educators. Examples are provided of items written at different cognitive levels, thereby avoiding tests that focus only on recall

and memorization of facts. The chapter also describes how to write questions about clinical practice or the nursing process and provides sample stems for use with those items. The types of items presented in the chapter are similar to those found on the NCLEX and many certification tests. When teachers incorporate these items on tests in nursing courses, students acquire experience with this type of testing as they progress through the program, preparing them for taking licensure and certification examinations as graduates.

Through papers and other written assignments, students develop an understanding of the content they are writing about. Written assignments with feedback from the teacher also help students improve their writing ability, an important outcome in any nursing program from the beginning level through graduate study. Chapter 9 provides guidelines for assessing formal papers and other written assignments in nursing courses. The chapter includes criteria for assessing the quality of papers, an example of a scoring rubric, and suggestions for assessing and grading written assignments.

Chapter 10 explains how to assemble, administer, and score a test. In addition to preparing a test blueprint and skillful construction of test items, the final appearance of the test and the way in which it is administered can affect the validity of its results. In Chapter 10, test design rules are described; suggestions for reproducing the test, maintaining test security, administering it, and preventing cheating are presented in this chapter as well.

Online education in nursing continues to expand at a rapid pace. Chapter 11 discusses assessment of learning in online courses, including testing and evaluating course assignments. The chapter begins with a discussion of online testing. To deter cheating and promote academic integrity, faculty members can use a variety of both low- and high-technology solutions. Providing timely and substantive feedback to students is critical in online courses, and we have included a sample rubric for an online discussion forum assignment and evaluation. Clinical evaluation of students in online courses and programs presents challenges to faculty members and program administrators. The chapter includes discussion of methods for evaluating students' clinical performance in an online course. Other sections of this chapter examine assessment of online courses, student evaluation of teaching, and evaluating the quality of online nursing programs.

After administering the test, the teacher needs to score it, interpret the results, and then use the results to make informed decisions. Chapter 12 discusses the processes of obtaining scores, performing test and item analysis, and interpreting the results (for both teacher-made and standardized tests). The chapter examines test score distributions, describes measures of central tendency and variability, and explains their use in interpreting test scores. How to interpret the difficulty index and discrimination index and analyze each distractor are described. Examples of item analyses are provided for true–false, matching, multiple-choice, and multiple-response

items. Exhibits in the chapter illustrate these analyses so readers understand how to analyze and interpret the performance of individual test items. It also suggests ways in which teachers can use posttest discussions to contribute to student learning and seek student feedback that can lead to test-item improvement. Teachers often debate the merits of adjusting test scores by eliminating items or adding points to compensate for real or perceived deficiencies in test construction or performance. We discuss this in the chapter and provide guidelines for faculty in making these decisions. A section of the chapter also presents suggestions and examples of developing a test-item bank. Many publishers also offer test-item banks that relate to the content contained in their textbooks; we discuss why faculty members need to be cautious about using these items for their own examinations.

Chapter 13 describes the process of clinical evaluation in nursing. It begins with a discussion of the outcomes of clinical practice in nursing programs and then presents essential concepts underlying clinical evaluation. In this chapter, we discuss fairness in evaluation, how to build feedback into the evaluation process, and how to determine *what* to evaluate in clinical courses.

Chapter 14 builds on concepts of clinical evaluation examined in the preceding chapter. Many evaluation methods are available for assessing competencies in clinical practice. We discuss observation and recording observations in notes about performance, checklists, and rating scales; written assignments useful for clinical evaluation such as journals, concept maps, case analyses, and short papers; electronic portfolio assessment and how to set up a portfolio system for clinical evaluation; conferences; and other methods such as group projects and self-evaluation. The chapter includes a sample form for evaluating student participation in clinical conferences and a rubric to use for peer evaluation of participation in group projects.

Simulation is used widely for instruction in nursing, and it also can be used for assessment. A simulation can be developed for students to demonstrate procedures and technologies, analyze data, and make decisions. Students can care for the patient individually or as a team. Student performance in these simulations can be assessed to provide feedback or to verify competencies. Some simulations incorporate standardized patients, actors who portray the role of a patient with a specific diagnosis or condition. Another method for evaluating skills and clinical competencies of nursing students is OSCE. In an OSCE, students rotate through stations where they complete an activity or perform a skill, which then can be evaluated. Chapter 15 describes these methods for assessing clinical competencies of students.

Chapter 16 explores social, ethical, and legal issues associated with testing and evaluation. Social issues such as test bias, grade inflation, effects of testing on self-esteem, and test anxiety are discussed. Ethical issues include privacy and access to test results. By understanding and applying codes for the responsible and ethical use of tests, teachers can ensure the proper use of assessment procedures and the valid interpretation of test results. We also discuss selected legal issues associated with testing.

Grading is the use of symbols, such as the letters A through F or pass–fail, to report student achievement. Grading is used for summative purposes, indicating how well the student met the outcomes of the course and clinical practicum. To represent valid judgments about student achievement, grades should be based on sound evaluation practices, reliable test results, and multiple assessment methods. Chapter 17 examines the uses of grades in nursing programs, types of grading systems, how to select a grading framework, and how to calculate grades with each of these frameworks. We also discuss grading clinical practice, using pass–fail and other systems for grading, and provide guidelines for the teacher to follow when students are on the verge of failing a clinical practicum. We also discuss learning contracts and provide an example of one.

Program evaluation is the process of judging the worth or value of an educational program. With the demand for high-quality programs, there has been a greater emphasis on systematic and ongoing program evaluation. Thus, Chapter 18 presents an overview of program evaluation models and discusses evaluation of selected program components, including curriculum, courses, and teaching–learning activities. The chapter includes content on accreditation, accrediting agencies in nursing, types of accreditation, and evaluation of distance education programs. We discuss the development of a systematic plan for evaluation and include a sample format. Student ratings of courses and teachers carry much weight in many schools of nursing; we discuss these ratings and related issues and examine other sources of information about teaching effectiveness.

In addition to this book, we have provided an Instructor's Manual that includes a sample course syllabus, chapter-based PowerPoint presentations, and ready-to-use modules for an online course (with chapter summaries, student learning activities, discussion forum questions, and assessment strategies).

We wish to acknowledge Adrianne Brigido, Director of Nurse Education for Springer Publishing Company, for her enthusiasm and continued support. We also thank Springer Publishing Company for its support of nursing education and for publishing our books for many years.

Marilyn H. Oermann
Kathleen B. Gaberson

INSTRUCTOR'S RESOURCES

Evaluation and Testing in Nursing Education, Sixth Edition, includes a robust ancillary package. Qualified instructors may obtain access to ancillary materials by emailing textbook@springerpub.com. Available resources:

- Instructor's Manual:
  - Sample Course Syllabus
  - Chapter Summaries
  - Student Learning Activities
  - Discussion Questions
  - Assessment Strategies
- Chapter-Based PowerPoint Presentations

CONCEPTS OF ASSESSMENT

1

ASSESSMENT AND THE EDUCATIONAL PROCESS

In all areas of nursing education and practice, assessment is important to obtain information about student learning, evaluate competencies and clinical performance, and arrive at other decisions about students and nurses. Assessment is integral to monitoring the quality of educational and healthcare programs. By evaluating outcomes achieved by students, graduates, and patients, the effectiveness of programs can be measured and decisions can be made about needed improvements.

Assessment provides a means of ensuring accountability for the quality of education and services provided. Nurses, like other healthcare professionals, are accountable to their patients and society in general for meeting patients' health needs. Along the same lines, nurse educators are accountable for the quality of teaching provided to learners, outcomes achieved, and overall effectiveness of educational programs. Educational institutions also are accountable to their governing bodies and society in terms of educating graduates for present and future roles. Through assessment, nurse educators and other healthcare professionals collect information for evaluating the quality of their teaching and programs as well as documenting outcomes for others to review. All educators, regardless of the setting, need to be knowledgeable about assessment, testing, measurement, and evaluation.

Assessment

Educational assessment involves collecting information to make decisions about learners, programs, and educational policies. Mislevy (2017) defined *assessment* as gathering information about what students know and can do. Are students learning the important concepts in the course and developing the clinical competencies? With information collected through assessment, the teacher can determine relevant learning activities to meet students' learning needs and help them improve performance. Assessment that provides information about learning needs is diagnostic; teachers use that information to decide on the appropriate content, learning activities, and practice opportunities for students to meet the desired learning outcomes.

Assessment also generates feedback for students, which is particularly important in clinical practice as students develop their competencies and learn to think through complex clinical situations. Feedback from assessment similarly informs the teacher and provides data for deciding how best to teach certain content and skills; in this way, assessment enables teachers to improve their educational practices and how they teach students.

Another important purpose of assessment is to provide valid and reliable data for determining students' grades. Although nurse educators continually assess students' progress in meeting the outcomes of learning and developing the clinical competencies, they also need to measure students' achievement in the course. Grades serve that purpose. Assessment strategies provide the data for faculty to determine whether students achieved the outcomes and developed the essential clinical competencies. Grades are symbols—for instance, the letters A through F—for reporting student achievement.

Assessment generates information for decisions about courses, the curriculum, and the nursing program. In this context, assessment is the process of collecting information for program evaluation and accreditation. Other uses of assessment information are to select students for admission to an educational institution and a nursing program and place students in appropriate courses. A broad view of assessment is that it encompasses the entire process of evaluating learners and institutional effectiveness (Banta & Palomba, 2014).

There are many assessment strategies that teachers can use to obtain information about students' learning and performance. These methods include tests that can be developed with different types of items, papers, other written assignments, projects, small-group activities, oral presentations, e-portfolios, observations of performance, simulation-based assessments, objective structured clinical examinations (OCSEs), and conferences, among others. Each of those assessment strategies as well as others is presented in this book.

Brookhart and Nitko (2019) identified five guidelines for effective assessment. These guidelines should be considered when deciding on the assessment strategy and its implementation in the classroom, online course, skills or simulation laboratory, or clinical setting.

1. *Identify the learning objectives (outcomes or competencies) to be assessed.* These provide the basis for the assessment: The teacher determines whether students are meeting or have met the outcomes and competencies. The clearer the teacher is about *what* to assess, the more effective will be the assessment.
2. *Match the assessment strategy to the learning goal.* The assessment strategy needs to provide information about the particular outcome or competency being assessed. If the outcome relates to analyzing issues in the care of patients with chronic pain, a true–false item about a pain medication would not be appropriate. An essay item, however, in which students analyze a

scenario about an adult with chronic pain and propose multiple approaches for pain management would provide relevant information for deciding whether students achieved that outcome.

3. *Meet the students' needs.* Students should be clear about what is expected of them. The assessment strategies, in turn, should provide feedback to students about their progress and achievement in demonstrating those expectations, and should guide the teacher in determining the instruction needed to improve performance.
4. *Use multiple assessment strategies and indicators of performance for each outcome.* It is unlikely that one assessment strategy will provide sufficient information about achievement of the outcomes. A test that contains mainly recall items will not provide information on students' ability to apply concepts to practice or analyze clinical situations. The extent and depth of student learning is often difficult to measure on a test. In most courses, multiple assessment strategies are needed to determine whether the outcomes were met.
5. *Keep in mind the limitations of assessment when interpreting the results.* One test, one paper, one observation in clinical practice, or one simulation activity may not be a true measure of the student's learning and performance. Many factors can influence the assessment, particularly in the clinical setting, and the information collected in the assessment is only a sample of the student's overall achievement and performance.

Tests

A test is a set of items to which students respond in written or oral form, typically during a fixed period of time. Brookhart and Nitko (2019) defined a *test* as an instrument or a systematic procedure for describing one or more characteristics of a student. Tests are typically scored based on the number or percentage of answers that are correct and are administered similarly to all students. Although students often dread tests, information from tests enables faculty to make important decisions about students.

Tests are used frequently as an assessment strategy. They can be used at the beginning of a course or instructional unit to determine whether students have the prerequisite knowledge for achieving the outcomes or whether they have already met them. With courses that are competency based, students can then progress to the next area of instruction. Test results also indicate gaps in learning and performance that should be addressed first. With that information, teachers can better plan their instruction. Tests can be used during the instruction to provide the basis for formative assessment (Miller, Linn, & Gronlund, 2013). This form of assessment is to monitor learning

progress, provide feedback to students, and suggest additional learning activities as needed. When teachers are working with large groups of students, it is difficult to gear the instruction to meet each student's needs. However, diagnostic quizzes and tests reveal content areas in which individual learners may lack knowledge. Not only do the test results guide the teacher in suggesting remedial learning activities, but they also serve as feedback to students about their learning needs. In some nursing programs, students take commercially available tests as they progress through the curriculum to identify gaps in their learning and prepare them for taking the National Council Licensure Examinations, the NCLEX-RN® or NCLEX-PN®.

Tests are used for selecting students for admission to higher education settings and to nursing programs. Admission tests provide norms that allow comparison of the applicant's performance with that of other applicants. Tests also may be used to place students into appropriate courses. Placement tests, taken after students have been admitted, provide data for determining which courses they should complete in their programs of study. For example, a diagnostic test of statistics may determine whether a nursing student is required to take a statistics course prior to beginning graduate study.

By reviewing test results, teachers can identify content areas that students learned and did not learn in a course. With this information, faculty can modify the instruction to better meet student learning needs in future courses. Last, testing may be an integral part of the curriculum and program evaluation in a nursing education program. Students may complete tests to measure program outcomes rather than to document what was learned in a course. Test results for this purpose often suggest areas of the curriculum for revision and may be used for accreditation reports.

Measurement

Measurement is the process of assigning numbers to represent student achievement or performance, for instance, answering 85 out of 100 items correctly on a test. The numbers or scores indicate the degree to which a learner possesses a certain characteristic. Measurement is important for reporting the achievement of learners on nursing and other tests, but not all outcomes important in nursing practice can be measured by testing. Many outcomes are evaluated qualitatively through other means, such as observations of performance in clinical practice or simulation.

Although measurement involves assigning numbers to reflect learning, these numbers in and of themselves have no meaning. Scoring 15 on a test means nothing unless it is referenced or compared with other students' scores or to a predetermined standard. Perhaps 15 was the highest or lowest score on the test, compared with other students. Or the student might have set a personal goal of achieving 15 on the test; thus, meeting this goal is more important than how others scored on the test.

Another interpretation is that a score of 15 might be the standard expected of this particular group of learners. To interpret the score and give it meaning, having a reference point with which to compare a particular test score is essential.

In clinical practice, how does a learner's performance compare with that of others in the group? Did the learner meet the outcomes of the clinical course and develop the essential competencies regardless of how other students in the group performed in clinical practice? Answers to these questions depend on the basis used for interpreting clinical performance, similar to interpreting test scores.

Norm-Referenced Interpretation

There are two main ways of interpreting test scores and other types of assessment results: norm referencing and criterion referencing. In norm-referenced interpretation, test scores and other assessment data are compared with those of a norm group. Norm-referenced interpretation compares a student's test scores with those of others in the class or with some other relevant group. The student's score may be described as below or above average or at a certain rank in the class. Problems with norm-referenced interpretations, for example, "grading on a curve," are that they do not indicate what the student can and cannot do, and the interpretation of a student's performance can vary widely depending on the particular comparison group selected.

In clinical settings, norm-referenced interpretations compare the student's clinical performance with the performance of a group of learners, indicating that the student has more or less clinical competence than others in the group. A clinical evaluation instrument in which student performance is rated on a scale of below to above average reflects a norm-referenced system. Again, norm-referenced clinical performance does not indicate whether a student has developed desired competencies, only whether a student performed better or worse than other students.

Criterion-Referenced Interpretation

Criterion-referenced interpretation, on the other hand, involves interpreting scores based on preset criteria, not in relation to the group of learners. With this type of measurement, an individual score is compared with a preset standard or criterion. The concern is how well the student performed and what the student can do regardless of the performance of other learners. Criterion-referenced interpretations may (a) describe the specific learning tasks a student can perform, for example, define medical terms; (b) indicate the percentage of tasks performed or items answered correctly, for example, define correctly 80% of the terms; and (c) compare performance against a set standard and decide whether the student met that standard, for example, met the medical terminology competency (Miller et al., 2013). Criterion-referenced interpretation determines how well the student

performed at the end of the instruction in comparison with the outcomes and competencies to be achieved.

With criterion-referenced clinical evaluation, student performance is compared against preset criteria. In some nursing courses, these criteria are the objectives or outcomes of the course to be met by students. In other courses, they are the competencies to be demonstrated in simulation or clinical practice, which are then used as the standards for evaluation. Rather than comparing the performance of the student with others in the group, and indicating that the student was above or below the average of the group, in criterion-referenced clinical evaluation, performance is measured against the outcomes or competencies to be demonstrated. The focus with criterion-referenced clinical evaluation is whether students achieved the outcomes of the course or demonstrated the essential clinical competencies, not how well they performed in comparison with the other students.

Evaluation

Evaluation is the process of making judgments about student learning and achievement, clinical performance, employee competence, and educational programs, based on the assessment data. In nursing education, evaluation typically takes the form of judging student attainment of the outcomes of the course and knowledge gained in it, and the quality of student performance in the clinical setting. With this evaluation, learning needs are identified, and additional instruction can be provided to assist students in their learning and in developing competencies for practice. Similarly, evaluation of employees provides information on their performance at varied points in time as a basis for judging their competence.

Evaluation extends beyond a test score or performance rating. Brookhart and Nitko (2019) defined *evaluation* as the process of making a value judgment about the worth or quality of a student's performance or of products developed by students representing their learning. With evaluation, the teacher makes value judgments about learners: *value* is part of the word *evaluation*. Questions, such as "How *well* did the student perform?" and "Is the student *competent* in clinical practice?" are answered by the evaluation process. The teacher collects and analyzes data about the student's performance, then makes a value judgment about the quality of that performance.

In terms of educational programs, evaluation includes collecting information *prior* to developing the program, *during* the process of program development to provide a basis for ongoing revision, and *after* implementing the program to determine its effectiveness. With program evaluation, faculty members collect data about their students, alumni, curriculum, and other dimensions of the program for the purposes of documenting the program outcomes, judging the quality of the program, and making sound decisions about curriculum revision. As educators measure outcomes for accreditation and evaluate their courses and curricula, they are engaging in program

evaluation. Although many of the concepts described in this book are applicable to program evaluation, the focus instead is on evaluating learners, including students in all types and levels of nursing programs and nurses in healthcare settings. The term *students* is used broadly to reflect both of these groups of learners.

Formative Evaluation

Evaluation fulfills two major roles: It is both formative and summative. Formative evaluation judges students' progress in meeting the desired outcomes and developing clinical competencies. With formative evaluation, the teacher judges the quality of the achievement while students are still in the process of learning (Brookhart & Nitko, 2019). Formative evaluation occurs throughout the instructional process and provides feedback for determining where further learning is needed.

With formative evaluation, the teacher assesses student learning and performance, gives students prompt and specific feedback about the knowledge and skills that still need to be acquired, and plans further instruction to enable students to fill their gaps in learning. Considering that formative evaluation is diagnostic, it typically is not graded. The purpose of formative evaluation is to determine where further learning is needed. In the classroom, formative information may be collected by teacher observation and questioning of students, diagnostic quizzes, small-group activities, written assignments, and other activities that students complete in and out of class. These same types of strategies can be used to assess student learning in online courses.

In clinical practice and other practice environments, such as simulation and skills laboratories, formative evaluation is an integral part of the instructional process. The teacher continually makes observations of students as they learn to provide patient care and develop their competencies, questions them about their understanding and decisions, discusses these observations and judgments with them, and guides them in how to improve performance. With formative evaluation, the teacher gives feedback to learners about their progress in achieving the outcomes of practice and how they can further develop their knowledge and competencies.

Summative Evaluation

Summative evaluation, on the other hand, is end-of-instruction evaluation designed to determine what the student has learned. With summative evaluation, the teacher judges the quality of the student's achievement in the course, not the progress of the learner in meeting the outcomes. Although formative evaluation occurs on a continual basis throughout the learning experience, summative evaluation is conducted on a periodic basis, for instance, every few weeks or at the midterm and final evaluation periods. This type of evaluation is "final" in nature and serves as a basis for grading and other high-stakes decisions.

Summative evaluation typically judges broader content areas and competencies than formative evaluation. Strategies used commonly for summative evaluation in the classroom and online courses are tests, papers, other assignments, and projects. In clinical practice, rating scales, written assignments, e-portfolios, projects completed about clinical experiences, and objective structured clinical examinations (OSCEs) may be used. Another strategy for summative evaluation is simulation, which can be used to assess students' decisions, skills, communication, teamwork, and other competencies.

Both formative and summative evaluation are essential components of most nursing courses. However, because formative evaluation represents feedback to learners with the goal of improving learning, it should be the major part of any nursing course. By providing feedback on a continual basis and linking that feedback with further instruction, the teacher can assist students in developing the knowledge and skills they lack.

Evaluation and Instruction

Figure 1.1 demonstrates the relationship between evaluation and instruction. The intended learning outcomes are the knowledge, skills, and competencies students are to achieve. Following assessment to determine gaps in learning and performance, the

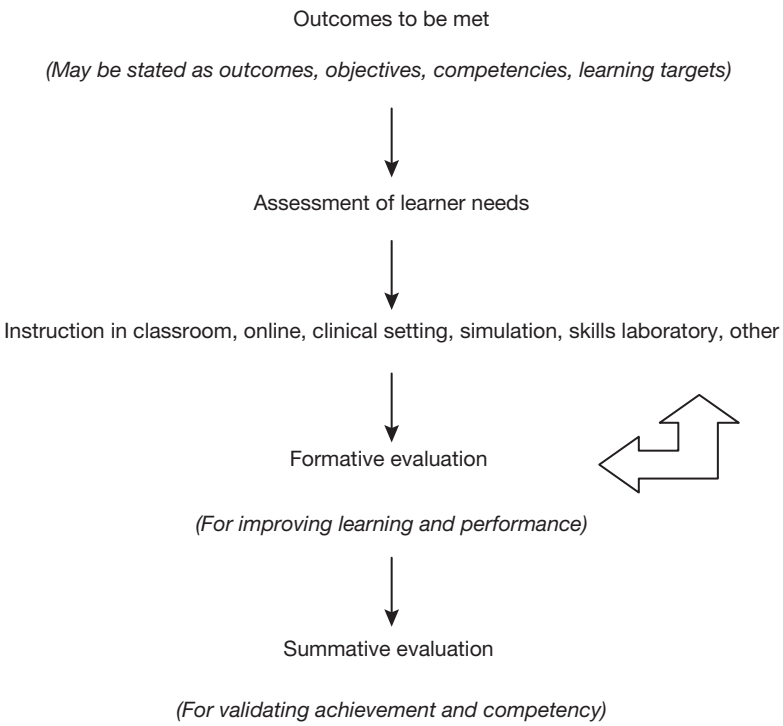

FIGURE 1.1 Relationship of evaluation and instruction.