

Chapter 7

A Standardized Clinical Evaluation Tool-Kit: Improving Nursing Education and Practice

Stephanie D. Holaday and Kathleen M. Buckley

A report released by the Institute of Medicine (IOM, 2003) informs us that “nurses and other health professionals are not being adequately prepared to provide the highest quality and safest care possible, and there is insufficient assessment of their proficiency” (p. 1). In response there has been much attention dedicated to curricular and teaching methodology reform, but unfortunately the development of sound clinical evaluation methods has often been an afterthought. This raises the question: In nursing education, is the dog wagging the tail or is the tail wagging the dog?

Serious attention needs to be devoted to reducing educational inconsistencies by linking the essential outcomes required for a degree in nursing with clinical evaluation outcomes in order for nurse educators to be accountable to their students, patients, and society (Holaday, 2004, 2006; Holaday, Buckley, & Miklancie, 2006). Although curricular requirements are similar in nursing schools throughout the United States, there is little consistency in expected outcomes that are assessed in the clinical setting and the methods used for evaluating them. Often, faculty rely on personal perceptions and instincts about expectations in the clinical setting, resulting in much variation. Variation in expected clinical outcomes leads to variation in actual clinical outcomes and proficiency. Therefore, the variations and inconsistencies of assessment

and evaluation in the clinical setting are strongly linked to the inadequate preparation of nurses in general.

Factors that partly contribute to the variations in clinical evaluation are the differences in clinical settings, patient populations, and experiences. However, the greatest contribution is associated with a lack of standardized clinical outcomes and instruments to measure those outcomes, confounded by the difficulty of measuring the subjective variables inherent in any clinical environment. Nevertheless, the need to evaluate the safety and competency of nursing students in the clinical setting remains, regardless of problems associated with doing so. If the assumption is that the performance that is assessed and measured in the clinical setting should tightly align with expected outcomes, then we must bridge the gap of variations and inconsistencies in the clinical setting with standardized outcomes, and strengthen the evaluation methods of those outcomes, to ensure that our nurse graduates are equipped with the knowledge and skills necessary to practice as competent health care professionals.

Nurse educators must ask the following questions: (a) How are the outcomes, deemed essential for safe and competent patient care by our professional organizations, incorporated into clinical assessment and evaluation of our nursing students? (b) What is the driving force behind instruction and evaluation of our nursing students: is it the outcomes recognized by the profession, or is it content of the nursing licensure examination, that is, “teaching to the test”? (c) How are nurse educators assessing and measuring student outcomes in the clinical setting? and (d) What evidence supports the evaluations given to our students in the clinical setting, that is, evidence-based teaching?

In this chapter, we present an innovative tool-kit used by three major university nursing schools to assess, evaluate, and measure student performance and growth across clinical settings and at all levels of educational preparation in a nursing program. The tool-kit includes consensus-based clinical outcome objectives and competencies against which to base judgments for evaluation, a five-point rating scale adapted from Bondy’s (1983) criterion-referenced matrix for measuring the quality of clinical performance, and conversion scales for grading the achievement of the clinical objectives, which are leveled according to educational preparation and correlate with the university grading scale.

The tool-kit is a blueprint for assessing and evaluating levels of clinical competence and can be used by any nationally accredited nursing school. It also provides a benchmark for assessing and measuring course and program outcomes. It is hoped that the guidance provided in this chapter will assist other nurse educators to customize a clinical evaluation instrument and method that is based on standardized outcomes, rooted in the values of the nursing profession, and suited to the individual needs of their institution.

Background

The Health Resources and Services Administration (HRSA) (2006) has projected that by the year 2020 the shortage of nurses in the United States will increase to more than 1 million, and if current trends continue, only 64% of this projected demand will be met. Even though the American Association of Colleges of Nursing (AACN, 2006, December 6) reported significant gains in both enrollments in nursing programs and graduates from those programs for 6 consecutive years, with a 5% increase in enrollment and an 18% increase in graduates in 2006, it's not enough. The HRSA (2006) report estimated that nursing schools must increase the number of graduates by 90% in order to adequately address the nursing shortage. Though interest in professional nursing careers seems strong, the AACN (2006, December 6) reported that because of a shortage in nursing faculty, colleges and universities turned away over 32,000 qualified applicants this past year, a trend that has increased over the past 6 years.

In answer to calls to accelerate the growth of the nursing workforce, nurse educators have been exploring creative strategies to boost enrollments and expand and accelerate nursing programs. Similarly, nurse educators are responding to the increasing amount and complexity of content that students need to learn in order to provide safe and effective nursing care, with innovative teaching–learning strategies and curriculum design. It is important, however, not to lose sight of the challenge inherent in efforts to produce nurses quickly as the knowledge explosion continues, and that is to ensure competency upon entry into practice.

The Institute of Medicine (IOM) (2000) estimated that as many as 98,000 patients die in hospitals each year from preventable medical errors, illustrating the importance of closely monitoring the progress of nursing students and ensuring a high degree of objectivity of their evaluation. A wide array of publications have cited the need to expand the ability of health care professionals to address concerns related to patient safety, performance improvement, and quality in health care (AACN, 2006). Lenburg (1999) noted that employers reported experiencing a widening gulf between the competencies required for practice and those that new graduates learned in their education programs. Employers are spending increasing amounts of time and resources to orient and teach new nurses the competencies required in today's workplace.

The National League for Nursing (2005) position statement emphasizes the need for nursing programs to align with the realities of the health care setting. Calls from authorities are loud and clear for health care disciplines to address the need for safe and competent care in our practice environments (Agency for Healthcare Research and Quality, 2003; IOM, 2004; Trust for America's Health, 2005). Through expert consensus, the AACN (2006) has identified specific educational competencies required for safe nursing practice, and it is revising the existing *Essentials of Baccalaureate Education for Professional Nursing Practice* (AACN, 1998) to reflect changes in the health care environment. It would seem to follow that nurse educators should link these educational competencies deemed essential by experts in the nursing profession to their curriculum and evaluation outcomes.

The need to have a certain standard of competence and expertise is common to all health professions and is important to protect the public, ensure quality of care, and establish credibility of the profession. In order to achieve standards of competence, Oermann and Gaberson (2006) highlight the following components in nursing education: (a) educational outcome objectives that reflect standardized competencies, (b) a valid and reliable method to determine if the educational objectives were achieved, and (c) a valid and reliable mechanism to determine if students can apply their knowledge toward implementing safe and effective patient care in real-life situations in the clinical setting. It is the last of these three components that has historically been problematic across the disciplines in health care education.

Although successful completion of the nursing licensure examination is a useful indicator of competency, it does not measure the overall competence of a graduate nurse. It does not provide sufficient evidence that graduate nurses have been prepared to apply their knowledge in providing safe patient care in professional practice. Standardized written exams are reliable in documenting recall, but they lack validity in assessing combined cognitive and psychomotor skills. They do not capture all the complexities and nuances of live patient encounters and the environments in which judgments and decisions are made. Therefore, they do not provide a complete picture of a student's performance by entry into practice.

Challenges With Clinical Performance Evaluation

Evaluating clinical performance is a daunting and complex task, and development of a valid and reliable instrument to do so is even more complex and fraught with problems. There are difficulties and frustrations associated with both the instrument and the user, so much so that a desired level of correlation between students' actual scores and scores on a criterion measure cannot be recommended for reliability because the correlation is influenced by many factors, including values, judgments, expertise, and experience of the evaluator (Oermann & Gaberson, 2006). These factors link the psychometric and social areas of research.

Clinical evaluations rely on observations, and these observations are highly subject to observer bias, resulting in little consistency among the scores of the evaluators and errors that threaten the quality of the data collected. Thompson, Lipkin, Gilbert, Guzzo, & Roberson (1990) found the halo effect, specifically the error of leniency, to be evident in evaluations of medical residents, reporting that 96% of the rating scores fell between 6 and 9 on a 10-point scale.

In examining the numerous evaluation tools available for use in the clinical setting, it appears that efforts to assist evaluators to rate students more objectively have resulted in outcomes and objectives that are more specifically stated, narrowing their use. Oermann and Gaberson (2006) explain that specific instructional objectives while reducing some of the subjectivity, limit the flexibility required in the clinical

setting and reduce the measurement focus to a more simple knowledge and skill level (Gronlund, 2000). It is clear that specific instructional objectives or competencies for clinical evaluation are inappropriate for use in nursing education, where critical thinking, judgments, and problem solving are vital requirements. Furthermore, when the nature of the clinical setting is unpredictable and varies with each experience, objectives must serve as a guide, allowing instructors to develop the appropriate instruction to elicit the outcome to be learned.

Accuracy and Reliability of Performance Ratings

Approaches to improve the accuracy and reliability of performance ratings have resulted in disappointment and frustration. Two major methods that have been studied to improve performance rating accuracy and reliability are (a) rater format training, focusing on the content of the rating scale, and (b) rater error training, focusing on elimination of common rating errors. Heneman (1988) concluded that neither of these rater training methods increased performance rating accuracy. Although error training has been shown to be successful in reducing psychometric errors, there has been substantial evidence that reducing psychometric error has little or no corrective effect on the accuracy, reliability, or discrimination of scores, because performance ratings are really a reflection of the skill of the raters (Gray, 1996; McIntyre & Smith, 1984). Furthermore, Heneman suggests that the generalizability of most findings on rater training is open to question.

While there are difficulties with rating scales, Oermann and Gaberson (2006) recommend using them for clinical evaluation because they allow instructors to rate performance over time and note patterns of performance. Unlike checklists that use two points on a scale to represent pass or fail results, a rating scale can be made up of three or more points, conveying far more information about the quality of the performance.

Methods to Improve Performance Ratings

Calman, Watson, Norman, Redfern, and Murrells (2002) propose that some of the difficulties with performance ratings can be addressed

through the development of a national instrument for standardizing clinical competence assessment and evaluation. Haber and Avins (1994) and Borman (1975) recommend standardizing the observation of behaviors and defining the nomenclature for the desired expectations in order to help reduce error biases resulting from the halo effect. Chirico (2004) found that when creating common rater performance expectations by providing raters with rating standards as well as examples (critical indicators) of these dimensions, idiosyncratic rating tendencies were reduced and accuracy was improved. In summary, a standardized assessment and evaluation instrument, based on clinical outcomes recognized by the profession, which seeks evidence-based knowledge of students in real-life situations, would support a more valid, reliable, and complete assessment and measurement of a student's overall competence.

Development of the Evaluation Tool-Kit

To address many of the problems associated with evaluating the clinical competence of nursing students, a clinical evaluation tool-kit was developed by one of the authors (Holaday, 2004). The tool-kit remains a work in progress. Since its inception, faculty from several university nursing schools, clinical educators and preceptors, statisticians, independent education consultants, and students have contributed to further development of the tool-kit. In addition, education staff from the AACN served as reviewers and consultants for the clinical outcome objectives and essential competencies in the tool-kit.

With permission of the developer (Holaday), the evaluation tool-kit was successfully piloted and implemented in all undergraduate clinical courses at the Catholic University of America (CUA) and George Mason University (GMU). Again with permission, Michigan State University (MSU) nursing faculty adopted the evaluation tool-kit for use in all of their baccalaureate clinical courses. The evaluation tool-kit has demonstrated that the measurement of the performance of standardized competencies in the clinical setting not only is possible but can render many positive results. It is important to note, however, that the education of the user is critical in ensuring the validity and reliability of the outcome performance data.

The most recent undertaking germane to the further development of the tool-kit is its licensing to other nursing schools and its adaptation to digital format for electronic distribution and use. Investigators involved in this project believe it offers a unique opportunity to establish an absolute standard of clinical performance in nursing education based on direct observation of clinical skills.

Purpose of the Evaluation Tool-Kit

The purpose of the evaluation tool-kit was to improve the accuracy and process of clinical evaluation, reduce discrepancies among evaluators, and reflect growth in students' performance as they progress through the nursing program. The goal was to provide a blueprint for clinical teaching and learning with benchmarks for evaluating levels of clinical competence. The focus was on developing one instrument with sufficient flexibility to accommodate various learning settings, experiences, and levels. While characteristics of the clinical setting were known to vary, the framework of the instrument, measurement scales, and the evaluation process are considered stable.

Conceptual and Measurement Framework

A criterion-referenced conceptual and measurement framework was selected to systematically guide the development of the evaluation tool-kit, as well as guide how the evaluation process was to be operationalized. A criterion-referenced evaluation involves comparing the student's clinical performance with predetermined criteria, and not with the performance of other students in the group (Waltz, Strickland, & Lenz, 2005). This framework provided a lens to view clinical education and student performance. It served as the basis for identifying standardized criteria and critical aspects of care to be learned and for designing a measurement system that would minimize outside influences by the environment or the rater. The incorporation of standardized essential criteria intended for universal use by all nursing programs facilitates the portability of the tool-kit across clinical settings, levels, and institutions, allowing for customization.

Method

The process of developing the tool-kit involved identifying the standardized criteria by which students would be evaluated. These criteria were then organized in the form of outcome objectives and essential competencies describing what the student is required to learn by the end of the nursing program. The process also involved the design of measurement tools, including a scale to rate the amount and quality of learned information, and scales to measure and grade the achievement of the objectives and competencies. The measurement scales were based on absolute standards without regard to the achievement of other students. This was done to create a fairer assessment and grading system, but it was understood that it would not eliminate the problems inherent in judging the quality of the clinical performance.

Outcome Objectives With Essential Competencies. Eleven broad outcome objectives were drawn from the two AACN documents: *The Essentials of Baccalaureate Education for Nursing Practice* (AACN, 1998) and *The Essential Clinical Resources for Nursing's Academic Mission* (AACN, 1999). Following an extensive literature review, the outcome objectives displayed in Table 7.1 were identified to be common to all courses in a baccalaureate program and were written to reflect the mission, philosophy, and values of the school of nursing. The 11 objectives were then incorporated into the curriculum to form both the terminal objectives for the baccalaureate program and the outcome objectives for clinical practice required for a baccalaureate degree in nursing. These objectives, as incorporated into the tool-kit and curriculum, provided a foundation that rested on established protocol, reflecting the strength of the nursing profession, grounded in institutional identity, which in turn adds to their validity.

The 11 outcome objectives were then operationally defined with essential competencies, also drawn from the two AACN *Essentials* documents (AACN, 1998, 1999). Table 7.2 presents examples of three outcome objectives with their associated essential competencies required by graduation. These essential competencies were written to impart clarity in meeting each outcome objective. They equip faculty with an enabling framework of relevant yet broad indicators to facilitate

TABLE 7.1 The 11 Outcome Objectives

-
1. Exhibits caring to facilitate spiritual, mental, and physical health.
 2. Shows self-awareness in pursuing learning opportunities to enhance professional development and delivery of nursing care.
 3. Expresses effective communication.
 4. Uses professional collaboration in the management and delivery of health care.
 5. Exhibits integrity, honesty, and accountability.
 6. Uses the teaching–learning process when providing health education.
 7. Acts as an advocate for the client and the health care profession.
 8. Shows awareness of, and sensitivity to, the values and mores of clients in ethical decision making.
 9. Demonstrates leadership skills.
 10. Uses critical thinking to promote holistic health.
 11. Performs technical skills in a competent and efficient manner.
-

planning their clinical courses. The qualitative description of each competency provides direction for faculty to identify the more specific performance criteria or critical behaviors unique to their clinical course. The level in which each competency is written, indicating the expected performance by the end of the nursing program, provides a benchmark, enabling faculty to adjust the expected level of performance to the level of their clinical course.

To provide for a complete evaluation of a student's clinical performance, the objectives and essential competencies were written to include the cognitive, affective, and psychomotor domains of learning. This was to assist faculty to move beyond just the evaluation of skills performed to a more global assessment and evaluation process, integrating all domains of learning, including features of clinical judgments and critical thinking.

In an effort to establish content validity of the outcome objectives and essential competencies, inferences were made about the representativeness of their content, first by panels of expert faculty from schools of nursing, and then by AACN educational staff. Each item and the overall content were reviewed for accuracy and completeness to ensure that the objectives and competencies represented the intent of

TABLE 7.2 Examples of Outcome Objectives (2, 3, and 10) With Their Essential Competencies

Objective 2: Shows Self-Awareness in Pursuing Learning Opportunities to Enhance Professional Development and Delivery of Nursing Care

- ___ a. Exhibits progressive socialization toward professional nurse status, observing and emulating nurse role models, and inculcating professional values (“engaged”).
- ___ b. Seeks learning opportunities and resources to develop competence.
- ___ c. Honestly and accurately evaluates personal performance.
- ___ d. Responds professionally to feedback or correction.
- ___ e. Makes changes to improve practice.

Objective 3: Expresses Effective Communication

- ___ a. Produces clear, relevant, organized, and thorough writing.
- ___ b. Exhibits legally accurate and appropriate documentation.
- ___ c. Recognizes and uses appropriate medical terminology and abbreviations.
- ___ d. Uses various forms of communication to increase understanding.
- ___ e. Uses clear and open expression in dialogue and is engaged with person or audience.
- ___ f. Elicits preferences and values from clients, clarifying understanding.
- ___ g. Exhibits professional and therapeutic body language.
- ___ h. Listens attentively and respectfully without interruption or disruption.
- ___ i. Maintains self-control and dignity and responds to situations professionally, without blame or aggressive behavior.

Objective 10: Uses Critical Thinking to Promote Holistic Health

- ___ a. Integrates theory and research-based knowledge from behavioral, biological, and natural sciences to analyze and interpret information.
 - ___ b. Gathers appropriate data for assessment.
 - ___ c. Identifies appropriate nursing diagnoses, goals, and outcome criteria.
 - ___ d. Recognizes pathological processes and problems when they arise and intervenes appropriately.
 - ___ e. Exhibits an accurate understanding of the expected effects and possible complications that could result from interventions.
 - ___ f. Makes appropriate judgments and sound decisions in the management of care based on a clear and accurate understanding of the rationale.
 - ___ g. Evaluates effectiveness of achieving outcomes and modifies appropriately.
 - ___ h. Integrates principles of primary health care in delivery of care.
 - ___ i. Uses research findings to enhance and improve clinical practice.
-

the essential competencies laid out in the two AACN documents. Suggestions, comments, and feedback resulted in adjustments and editorial changes.

The Measurement Scales

The Rating Scale. In keeping with the criterion-referenced framework, it was decided that the rating scale for measuring the quality of performance with respect to each essential competency and outcome objective would be adapted from Bondy's criterion-referenced matrix (1983). Bondy's development of this five-point scale and its established validity and reliability have contributed greatly to nursing education and were major factors in the decision to use the scale. Modifications to some of Bondy's criterion-referenced label descriptors and categories were necessary in order to use the scale to rate performance of the same outcome objectives and essential competencies at all levels of clinical preparation. These modifications were accomplished through a faculty workgroup that incorporated comments and suggestions from both student and faculty focus groups from two university nursing schools. The resulting scale measures the amount of guidance required to accurately and efficiently perform each clinical competency. The modifications to Bondy's scale (1983) are described subsequently, and the resulting rating scale is displayed in Table 7.3.

The number of reference points on the scale remained at five (0–4). The label descriptor for scale point 4 was changed from *independent* to *self-directed*, and the descriptor for scale point 1 was changed from *marginal* to *novice*. These changes reflected concerns from both faculty and students regarding clarity of the terms *independent* and *marginal* on Bondy's scale. It was decided that the term *self-directed* better described the intent of scale point 4 because a nursing student cannot care for patients independently, and the term *novice* better described the intent of scale point 1 because the term *marginal* indicated a student who was barely performing within the lower standard or limit of quality. This would have presented a problem when using the scale in beginning clinical courses, where students are expected to perform at the novice

TABLE 7.3 The Rating Scale

Self-Directed (4)	
<p>Almost Never Requires (<10% of the time)</p> <ul style="list-style-type: none"> • direction • guidance • monitoring • support 	<p>Almost Always Exhibits (>90% of the time)</p> <ul style="list-style-type: none"> • a focus on the client or system • accuracy, safety & skillfulness • assertiveness and initiative • efficiency and organization • an eagerness to learn
Supervised (3)	
<p>Occasionally Requires (25% of the time)</p> <ul style="list-style-type: none"> • direction • guidance • monitoring • support 	<p>Very Often Exhibits (75% of the time)</p> <ul style="list-style-type: none"> • a focus on the client or system • accuracy, safety & skillfulness • assertiveness and initiative • efficiency and organization • an eagerness to learn
Assisted (2)	
<p>Often Requires (50% of the time)</p> <ul style="list-style-type: none"> • direction • guidance • monitoring • support 	<p>Often Exhibits (50% of the time)</p> <ul style="list-style-type: none"> • a focus on the client or system • accuracy, safety & skillfulness • assertiveness and initiative • efficiency and organization • an eagerness to learn
Novice (1)	
<p>Very Often Requires (75% of the time)</p> <ul style="list-style-type: none"> • direction • guidance • monitoring • support 	<p>Occasionally Exhibits (25% of the time)</p> <ul style="list-style-type: none"> • a focus on the client or system • accuracy, safety & skillfulness • assertiveness and initiative • efficiency and organization • an eagerness to learn

(continued)

TABLE 7.3 The Rating Scale (*continued*)

Dependent (0)	
Almost Always Requires (>90% of the time)	Almost Never Exhibits (<10% of the time)
<ul style="list-style-type: none"> • direction • guidance • monitoring • support 	<ul style="list-style-type: none"> • a focus on the client or system • accuracy, safety & skillfulness • assertiveness and initiative • efficiency and organization • an eagerness to learn

© Copyright 2004 by Stephanie Holaday. Reprinted with permission.

level, and use of the term *marginal* would have indicated a poor performance rather than an expected level of performance.

Bondy used three categories to describe criteria for clinical evaluation: (a) the safety and accuracy of the performance, (b) the proficiency, efficiency, and skillfulness of the performance, and (c) cues needed in order to perform a competency. These three categories were modified to two, and frequency descriptors were used along with the qualitative descriptors. The modified categories describe (a) the quality of the performance and (b) the amount of guidance required to perform the competency. The first category combines Bondy's first two categories. The second category reflects the amount of guidance required to perform the competency. Both categories use frequency labels to describe the amount of guidance required to perform a competency accurately, safely, and efficiently.

As students improve in accuracy, safety, and efficiency, they demonstrate a decrease in the amount of guidance required when performing the competency. Commensurately, as students progress from dependent to novice to supervised, and to self-directed, their scores should demonstrate an increase from 1 to 4. This reflects the work by Benner (1984), who defines levels of practice from the beginner as a novice to the expert who has developed a deep understanding and an intuitive grasp of situations.

The modifications to Bondy's scale illustrate the expected decreases in guidance as students gain knowledge and competence from requiring

90% or more guidance (almost always) to perform a competency, to requiring only 10% or less guidance (almost never) to perform a competency. Levels of expected performance that correspond to the rating scale are outlined in Table 7.4. This important feature highlights students' growth over time, as they progress through their nursing program, and provides a clear picture for student, course, and program evaluation.

The Rating Format. The evaluation format used to rate the quality of student performance of the outcome objectives and competencies (see Table 7.5) was adapted from the clinical evaluation tool developed and tested by Krichbaum, Rowan, Duckett, Ryden, and Savik (1994). The work accomplished by Krichbaum and colleagues (1994) and continued by Duckett and colleagues (1997) contributed significantly to the overall format and process used in this evaluation tool-kit, as well as to its validity. Krichbaum and investigators (1994) combined the use of Bondy's rating scale with broad clinical outcome measures to evaluate performance across clinical settings. Their findings included strong positive correlations between scores on their evaluation tool and variables such as grade point average, college credits earned, and moral reasoning. Additionally, the high Cronbach's alphas indicated that their evaluation tool was a reliable measure of clinical performance (Krichbaum et al., 1994).

The evaluation tool-kit builds on the work by Krichbaum and colleagues (1994) and Duckett and colleagues (1997) by incorporating a scale that provides meaning to the final rating scores at five clinical levels in the nursing program.

TABLE 7.4 Expected Levels of Performance

Level	Expected Performance
Performance I	Novice-Assisted
Performance II	Assisted
Performance III	Assisted-Supervised
Performance IV	Supervised-Self-directed
Performance V	Self-directed

TABLE 7.5 The Clinical Performance Rating Form

1. Exhibits caring to facilitate spiritual, mental and physical health:		
Self-directed	(4)	Objective 1
Supervised	(3)	___ Rating Score
Assisted	(2)	
Novice	(1)	
Dependent	(0)	
2. Shows self-awareness in pursuing learning opportunities to enhance professional development and delivery of nursing care:		
Self-directed	(4)	Objective 2
Supervised	(3)	___ Rating Score
Assisted	(2)	
Novice	(1)	
Dependent	(0)	
3. Expresses effective communication:		
Self-directed	(4)	Objective 3
Supervised	(3)	___ Rating Score
Assisted	(2)	
Novice	(1)	
Dependent	(0)	
4. Uses professional collaboration in the management and delivery of health care:		
Self-directed	(4)	Objective 4
Supervised	(3)	___ Rating Score
Assisted	(2)	
Novice	(1)	
Dependent	(0)	
5. Exhibits integrity, honesty, accountability:		
Self-directed	(4)	Objective 5
Supervised	(3)	___ Rating Score
Assisted	(2)	
Novice	(1)	
Dependent	(0)	

(continued)

TABLE 7.5 (continued)

6. Uses the teaching-learning process when providing health education:

Self-directed	(4)	Objective 6
Supervised	(3)	___ Rating Score
Assisted	(2)	
Novice	(1)	
Dependent	(0)	

7. Acts as an advocate for the client and health care profession:

Self-directed	(4)	Objective 7
Supervised	(3)	___ Rating Score
Assisted	(2)	
Novice	(1)	
Dependent	(0)	

8. Shows awareness of, and sensitivity to values and mores of clients in ethical decision making:

Self-directed	(4)	Objective 8
Supervised	(3)	___ Rating Score
Assisted	(2)	
Novice	(1)	
Dependent	(0)	

9. Demonstrates leadership skills:

Self-directed	(4)	Objective 9
Supervised	(3)	___ Rating Score
Assisted	(2)	
Novice	(1)	
Dependent	(0)	

10. Uses critical thinking to promote holistic health:

Self-directed	(4)	Objective 10
Supervised	(3)	___ Rating Score
Assisted	(2)	
Novice	(1)	
Dependent	(0)	

(continued)

TABLE 7.5. The Clinical Performance Rating Form (continued)

11. Performs technical skills in a competent and efficient manner:		
Self-directed	(4)	Objective 11
Supervised	(3)	___ Rating Score
Assisted	(2)	
Novice	(1)	
Dependent	(0)	
Total Rating Score for Clinical Performance		

© Copyright 2004 by Stephanie Holaday. Reprinted with permission.

The Grade Point Conversion Scale by Clinical Performance Level. In order to use the same rating scale to rate the same essential competencies and outcome objectives in all clinical courses but at five different levels of performance, a workgroup of seasoned clinical faculty developed a scale that converts the final score given on the rating scale to a grade point at five clinical performance levels. The scale converts the sum of the 11 rating scores given to the outcome objectives to a grade point at the end of each level of clinical preparation. These benchmarks indicate the competency level required to progress to the next level. Students move through the nursing program as they achieve the expected competency level of all essential outcome objectives.

The rating scale, used along with the grade point conversion scale, has been an excellent template for communicating faculty expectations and providing feedback to students for both formative and summative evaluation.

Each of the five levels on the grade point conversion scale represents a clinical course progression in the nursing program. Through a brainstorming technique, discussion, and consensus, the workgroup identified behavioral examples that represented a performance rating score of 4 (an A student) at the fifth level (senior) for each essential competency under each of the 11 outcome objectives. Then, using the rating scale as a guide, faculty teaching at performance levels one, two, three, or four assigned ratings to each behavioral example for the amount of guidance expected at their course level for an A student.

The ratings assigned to each behavioral example were averaged under each of the 11 outcome objectives. This provided a mean score for an A student for each of the 11 outcome objectives at performance levels one, two, three, and four. These 11 mean scores were then totaled for each of the five levels of performance, arriving at a numerical rating score that represented the uppermost end of the conversion scale for each of the five performance levels. A grade point of 4.0 was assigned to each of these ratings. From these anchor points, subsequent grade points were calculated for each level of performance to correlate with a university's grading scale.

The grade point conversion scale presented in Table 7.6 reflects two of the five levels of clinical practice (performance level I and performance level III) through which nursing students progress during their educational program according to one university's grading scale. It illustrates what Benner recognized as levels of growth over time, as the learner gains expertise. Grade point conversion scales were developed to correlate with each university's grading scale, and piloted along with the outcome objectives and the rating scale at five levels of clinical preparation in two university nursing schools (CUA and GMU). Slight adjustments were made to the top end of the grade point conversion scale for performance levels one, three, and four, followed by recalculating the subsequent grade points to correlate with each university's grading scale until there was consensus among faculty from both schools as to the accuracy of the meaning of the grade points on the scales.

TABLE 7.6 Example of Two Clinical Performance Levels Within a Grade Point Conversion Scale

Clinical Performance I							
F (0.0)	D (1.0)	C- (1.5)	C (2.0)	B- (2.5)	B (3.0)	A- (3.5)	A (4.0)
< 11	11–12	13	14–15	16	17–18	19–20	≥ 21
Clinical Performance III							
F (0.0)	D (1.0)	C- (1.5)	C (2.0)	B- (2.5)	B (3.0)	A- (3.5)	A (4.0)
< 21	21–23	24–25	26–27	28–29	30–31	32–33	≥ 34

The process that was undertaken to level the scores on the rating scale in order to develop the grade point conversion scale at five levels of clinical preparation draws on the value of faculty agreement, which was used successfully by Bondy when establishing validity and reliability of her criterion-referenced matrix scale. It also places value on our professional nurses' expert opinion, supporting much of Benner's work (1984) and recognizing how expert nurses are able to account for the many factors that make up real-life situations in clinical practice. If areas of agreement evolve between experts with substantial experience in clinical teaching, then the identification of rating scores that represent the expected performance of elements related to that teaching contributes to the validity and reliability of this clinical evaluation tool-kit.

The Evaluation Process

As previously discussed, this tool-kit provides options to faculty when evaluating clinical practice. It is important, however, to maintain as much consistency as possible between and among clinical courses. The following discussion outlines the basic steps of the evaluation process and then presents some of the options and alternatives offered by the evaluation tool-kit.

After decisions are made regarding the type of grading system—for example, pass-fail, 5 through 1, or A through F—the next consideration relates to what to include in the clinical course grade and the framework for determining that grade. In other words, if the clinical performance is graded, will it be weighted as 100% of the course grade, or will there be other components to be factored into the final grade, such as written assignments, lab work, demonstrations, and so forth? If other components are to be factored into the final grade, how much will each component be weighed.

Basic Steps of the Evaluation Process

The following steps summarize the basic evaluation process according to the evaluation tool-kit: (a) Each of the essential competencies

categorized under each of the 11 outcome objectives are rated (0–4); (b) a mean score of essential competencies is calculated for each of the 11 objectives; (c) the 11 mean scores given to the outcome objectives are totaled; and (d) the totaled score is converted to a grade point using the grade point conversion scale (see Table 7.6).

Arriving at the Final Grade

If grades are reported according to the grade point system, then the grade point calculated in the final step, described in the previous section, would reflect the final grade for the clinical course. If grades are reported according to a letter grade system, they should be easily converted to the final letter grade because the grade point conversion scale was based on the university grading scale.

If using a pass-fail grading system, one option might be to require all essential competencies and outcome objectives to be met at the acceptable performance level outlined in Table 7.4 in order to pass the course. If pass-fail is used for evaluating performance, other components may or may not be graded separately as part of the course grade. For example:

Written assignments	weighed 40%
Demonstration assignments	weighed 60%
Clinical performance rating	pass required

Alternative Scoring Considerations

If calculating each essential competency under the 11 categories is seen as cumbersome, an alternative would be to use the competencies as a guide for direction and planning and to only rate each individual competency. Other scoring decisions should be considered and agreed upon prior to implementing the evaluation, such as whether to round up or down. If rounding up is decided, will it be carried through all the calculations? If not, what is the point at which the score is rounded down? Also, if rounding up, from which decimal point is the score rounded?

Factoring in Other Components of the Clinical Course

Measurement experts caution against using only one method for determining the entire clinical course grade. If the aim is that our assessment reveals whether a student has achieved the complex knowledge and skills required for competent practice, different methods of competence assessment are recommended to be included as part of a final clinical grade. Different methods address different abilities, providing a more valid assessment and complete picture of student outcomes.

Strategies such as graded demonstrations or written assignments can easily be factored into the final course grade to support the validity of inferences made on clinical performance. For example, if assignments are graded according to a percent scale, the final assignment percent grade may be converted to a grade point using the university's conversion scale. This results in a grade point for assignments and a grade point for the clinical performance evaluation. These two grade points can be weighed differently or equally for the final course grade.

Discussion

The teaching–learning process has not been completed until information required to be learned is assessed and evaluated for subsequent use. This is a necessary process in nursing education. Regardless of the grading system or strategies used for clinical evaluation, this tool-kit provides a blueprint for collecting the data needed and determining if students' performance reflects achievement of the competencies required for clinical practice. Although there are program, curriculum, and grading system differences across nursing education, this evaluation tool-kit is applicable for use in varied clinical courses, settings, and educational levels.

Limitations

Although some faculty have reported difficulty using the percent frequency labels on the rating scale, other faculty appreciate them. Because

there seems to be a split belief as to their value, the percent frequency labels were left as part of the rating scale, and the individual course or program faculty can decide whether to use them. It should be noted that Kroboth and coworkers (1992) concluded that performance of a clinical behavior has to be observed 6–10 times to achieve a reliability coefficient of 0.8. Limited exposure to encounters in the clinical setting for direct observation and evaluation of a given experience compromises any assessment reliability. This is known as an assessment gap, which is the difference between the amount of evidence collected and the amount of evidence needed to make an inference about the quality of a performance. It could be argued that the use of the percent frequency labels compels clinical faculty to actively seek opportunities for observing and assessing behaviors for evaluation, which in turn provides for accountability.

In the face of the many variations in clinical environments and patient situations, and the problems associated with rater differences, the measurement format in this tool-kit provides general guidance, and the performance indicators to be measured are general in their capability to be measured. The final responsibility for adequately measuring student outcomes lies with the individual nursing faculty (the user). It is at this point that trust is placed in our nurse educators to make the proper interpretations and use of the evaluation instruments that evolve. This might be considered a limitation, but we believe that this presents an opportunity for faculty improvement through active involvement in the evaluation process rather than through passively receiving instructions and guidance. A balance between the inherent limitations and opportunities must be attained. Indeed, the most challenging issue to be dealt with, especially with the increasing distribution of the tool-kit, is to ensure that the concepts that are necessary for its proper use have been accurately communicated to the users. Training, collaboration, and communication across users will lead to an objective and value-free measurement tool and our common goal of performance improvement.

Implications

The evaluation tool-kit is a blueprint for evaluating levels of clinical competence and an investment to assure clinical agencies that

student graduates were prepared by standardized competencies at an accepted and safe level of competence. It will assist nurse educators to provide students with the support and guidance needed to focus their attention on expectations and critical behaviors to be performed, while demonstrating growth in their clinical competence throughout their nursing program (Holaday, et al., 2006). This should result in broader and deeper student learning and greater student success. When nursing students have a clear understanding of expectations and are evaluated by clear guidelines for levels of expected performance, an improvement in performance as well as the quality of care given by the time of entry into professional practice should result.

The evaluation tool-kit is also a blueprint for consistency in clinical education, with a built-in mechanism for accountability by the user through the assessment and evaluation process. It provides momentum for institutional change and supports continuous improvement, which leads to improvement across health care institutions. This work also has implications for professional nursing practice as a blueprint for the use of professional practice standards to evaluate nurses for clinical levels or ladders and promotion.

Conclusion

The clinical evaluation tool-kit presented in this chapter is a blueprint for preparing students for “readiness for safe, quality nursing practice.” It was designed to assist with a fair and valid clinical assessment and evaluation across the broad spectrum of nursing roles, responsibilities, settings, and educational levels. The collegiality and shared vision of faculty from the three universities currently using this tool-kit have generated energy, excitement, and a strong commitment to demand excellence in teaching, learning, assessment, and evaluation. At the time of this writing, a number of other nursing schools with both 4-year and 2-year programs have adopted the evaluation tool-kit as their blueprint for customizing an evaluation instrument and method based on the individual needs of their institutions. While studies are currently under way to assess the effectiveness and measure the outcomes of this innovative evaluation tool-kit,

investigators believe it offers a unique opportunity to establish an absolute standard of clinical performance in nursing education.

The AACN (1997) position statement discusses the importance of preparing nursing students' curricula in programs that are based on core nursing values. These values are well articulated in the two AACN documents used to develop this tool-kit. They are intended to serve as faculty expectations for student outcomes. Since a team of nursing experts developed the outcome indicators and essential competencies included in these two documents, it should follow that we use them as our gold standard for education, assessment, and evaluation of our nursing students. Moreover, these same indicators and competencies are the outcome criteria used by the AACN for the accreditation process. This process would flow smoothly if nurse educators used these outcomes for evaluating programs as well as courses, both theory and clinical, which in turn would direct the teaching of those courses within the overall nursing program. There would then be no question that "the dog is wagging the tail."

REFERENCES

- Agency for Healthcare Research and Quality. (2003). *Patient safety initiative: Building foundations, reducing risk*. Interim Report to the Senate Committee on Appropriations (AHRQ Publication No. 04-RG005). Rockville, MD: Author.
- American Association of Colleges of Nursing. (1997). *A vision of baccalaureate and graduate nursing education: The next decade*. Washington, DC: Author.
- American Association of Colleges of Nursing. (1998). *The essentials of baccalaureate education for professional nursing practice*. Washington, DC: Author.
- American Association of Colleges of Nursing. (1999). *The essential clinical resources for nursing's academic mission*. Washington, DC: Author.
- American Association of Colleges of Nursing. (2006). *Hallmarks of quality and patient safety: Recommended baccalaureate competencies and curricular guidelines to assure high quality and safe patient care*. Retrieved February 12, 2007, from <http://www.aacn.nche.edu/Education/pdf/PSHallmarks9-06.pdf>
- American Association of Colleges of Nursing. (2006, December 6). *Student enrollment rises in U.S. nursing colleges and universities for the 6th consecutive year*. American Association of Colleges of Nursing Press Release. Retrieved February 12, 2007, from <http://www.aacn.nche.edu/Media/NewsReleases/06Survey.htm>
- Benner, P. (1984). *From novice to expert: Excellence and power in clinical nursing practice*. Menlo Park, CA: Addison Wesley.
- Bondy, K. (1983). Criterion-referenced definitions for rating scales in clinical evaluation. *Journal of Nursing Education*, 22(9), 376-382.

- Borman, W. (1975). Effects of instructions to avoid halo error on reliability and validity of performance evaluation ratings. *Journal of Applied Psychology*, 60, 556–560.
- Calman, L., Watson, R., Norman, I., Redfern, S., & Murrells, T. (2002). Assessing practice of student nurses: Methods, preparation of assessors and student views. *Journal of Advanced Nursing*, 38(5), 516–523.
- Chirico, K. (2004). A note on the need for true scores in frame-of-reference (FOR) training research. *Journal of Managerial Issues*, XIV(3), 382–395.
- Duckett, L., Rowan, M., Ryden, M., Krichbaum, K., Miller, M., Wainwright, H., et al. (1997). Progress in the moral reasoning of baccalaureate nursing students between program entry and exit. *Nursing Research*, 46(4), 222–229.
- Gray, J. (1996). Global rating scales in residency education. *Academic Medicine*, 71(1), 55–63.
- Gronlund, N. (2000). *How to write and use instructional objectives* (6th ed.). Upper Saddle River, NJ: Prentice Hall.
- Haber, R., & Avins, A. (1994). Do ratings on the American Board of Internal Medicine Resident Evaluation Form detect differences in clinical competence? *Journal of General Internal Medicine*, 9(3), 140–145.
- Health Resources and Services Administration. (2006). *What is behind HRSA's projected supply, demand, and shortage of registered nurses?* National Center for Workforce Analysis Reports. Rockville, MD: Author. Retrieved February 12, 2007, from <http://bhpr.hrsa.gov/healthworkforce/reports/behindrnprojections/index.htm>
- Heneman, R. L. (1988). Traits, behaviors, and rater training: Some unexpected results. *Human Performance*, 1(2), 85–98.
- Holaday, S. (2004, November). *The clinical performance-based evaluation tool-kit: Standardizing nursing education*. Paper presented at the annual meeting of American Association of Colleges of Nursing, The 2004 Baccalaureate Education Conference, San Antonio, TX.
- Holaday, S. (2006, November). *A tool-kit for advancing clinical performance evaluation*. Paper presented at the annual meeting of American Association of Colleges of Nursing, The 2006 Baccalaureate Education Conference, Orlando, FL.
- Holaday, S., Buckley, K., & Miklancie, M. (2006, August). *Standardizing clinical performance evaluation through a clinical evaluation tool-kit and model*. Paper presented at the meeting of the Institute for Educators in Nursing and Health Professions on Educating Nurses: Innovation to Application, Baltimore, MD.
- Institute of Medicine. (2000). *To err is human: Building a safer health system*. Washington, DC: National Academies Press.
- Institute of Medicine. (2003). *Health professions education: A bridge to quality*. Washington, DC: National Academies Press.
- Institute of Medicine. (2004). *Keeping patients safe: Transforming the work environment of nurses*. Washington, DC: National Academies Press.
- Krichbaum, K., Rowan, M., Duckett, L., Ryden, M., & Savik, K. (1994). The clinical evaluation tool: A measure of the quality of clinical performance of baccalaureate nursing students. *Journal of Nursing Education*, 33(9), 395–403.

- Kroboth, F., Hanusa, B., Parker, S., Coulehan, J., Kapoor, W., & Brown, F. (1992). The inter-rater reliability and internal consistency of a clinical evaluation exercise. *Journal of General Internal Medicine*, 7(2), 174–179.
- Lenburg, C. (1999, September 30). Redesigning expectations for initial and continuing competence for contemporary nursing practice. *Online Journal of Issues in Nursing*, 4(2). Retrieved February 12, 2007, from http://www.nursingworld.org/ojin/topic10/tpc10_1.htm
- McIntyre, R., & Smith, D. (1984). Accuracy of performance ratings as affected by rater training and perceived purpose of rating. *Journal of Applied Psychology*, 69(1), 147–156.
- National League for Nursing, Board of Governors. (2005). Position statement: Transforming nursing education. *Nursing Education Perspectives*, 26(3), 195–197.
- Oermann, M. H., & Gaberson, K. (2006). *Evaluation and testing in nursing education* (2nd ed.). New York: Springer.
- Thompson, W., Lipkin, M., Gilbert, D., Guzzo, R., & Roberson, L. (1990). Evaluating evaluation: Assessment of the American Board of Internal Medicine Resident Evaluation Form. *Journal of General Internal Medicine*, 5(6), 214–217.
- Trust for America's Health. (2005). *Ready or not? Protecting the public's health from diseases, disasters and bioterrorism*. Washington, DC: Author. Retrieved February 12, 2007, from <http://www.rwjf.org/files/publications/other/TFAH2006Revised.pdf>
- Waltz, C., Strickland, O., & Lenz, E. (2005). *Measurement in nursing and health research*. New York: Springer.

Contributors

Rebecca Bofinger, MSN, RN

Instructor

The Breen School of Nursing
Ursuline College
Pepper Pike, Ohio

Wanda Bonnel, PhD, RN

Associate Professor

University of Kansas School of
Nursing
Kansas City, Kansas

Kathleen M. Buckley, PhD, RN

Associate Professor

School of Nursing
The Catholic University of
America
Washington, DC

Susan E. Campbell, PhD, RN

Professor

Department of Nursing
College of St. Catherine
St. Paul, Minnesota

W. Dean Care, EdD, RN

Associate Professor and Interim
Dean

Faculty of Nursing
University of Manitoba
Winnipeg, Manitoba, Canada

**Peggy L. Chinn, PhD, RN,
FAAN**

Professor Emerita of Nursing
University of Connecticut
Editor, *Advances in Nursing
Science*
Emeryville, California

Jan Coye, PhD, RN

Academic Community
Liaison

Kirkhof College of Nursing
Grand Valley State University
Grand Rapids, Michigan

Jean Davis, EdD, ARNP, BC

Adjunct Faculty
Nova Southeastern University
Fort Lauderdale, Florida

Debra A. Filer, PhD, RN

Associate Professor
Department of Nursing
College of St. Catherine
St. Paul, Minnesota

Maria Fletcher, PhD, RN

Associate Professor
St. Joseph's College
Brooklyn, New York

Betsy Frank, PhD, RN

Professor
Indiana State University
College of Nursing Health
and Human Services
Terre Haute, Indiana

**Virginia Gilbert, MS, IBCLC,
RNC**

Assistant Professor
Department of Nursing
Boise State University
Boise, Idaho

**Nel Glass, PhD, RN,
FCN(NSW), FRCNA**

Associate Professor
Department of Nursing and
Health Care Practices
School of Health and Human
Sciences
Southern Cross University
Lismore, NSW, Australia

Janice J. Hoffman, PhD, RN

Instructor
Johns Hopkins University
School of Nursing
Baltimore, Maryland

Janice Hodges, MSN, RN

Director, Nursing Practice,
Education and Research
Sparrow Health System
Lansing, Michigan

Stephanie D. Holaday, DrPH, RN

Assistant Professor
College of Health & Human Services

School of Nursing
George Mason University
Fairfax, Virginia

Kay Setter Kline, PhD, RN

Professor
Kirkhof College of Nursing
Grand Valley State University
Grand Rapids, Michigan

Charlene Ludwig, MS, ARNP-BC

Assistant Professor
Johnson County Community
College
Overland Park, Kansas

**Gayle Preheim, EdD, RN,
CNAA, BC, CNE**

Director, Baccalaureate Nursing
Program
Associate Professor
School of Nursing
University of Colorado at
Denver and Health Sciences
Center
Denver, Colorado

Karen Rizk, MSN, RN

Instructor
The Breen School of Nursing
Ursuline College
Pepper Pike, Ohio

Judith M. Scanlan, PhD, RN

Associate Professor and
Associate Dean, Graduate
Program
Faculty of Nursing
University of Manitoba
Winnipeg, Manitoba, Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.