Evaluation of student performance in the clinical environment is one of the most challenging functions of the clinical instructor. In the first of a two part education feature, the author defines evaluation and describes the purpose, role, focus, research and legal implications of clinical evaluation, as well as types of evaluation instruments and procedures.

Education is a process of change. The cognitive, affective and psychomotor domains are the essence of learning and must be addressed in the evaluation process. Evaluation, in a general sense, is the process of describing some quality or characteristic of an individual, a program or an institution as the basis for making a judgment about that individual, program or institution. Lynch states that "the evaluation of educational outcomes remains one of the stubbornest problems facing education today.”

The task of evaluation is complicated by the multiplicity of variables involved in this dynamic, interactive process.

Several curriculum theorists have developed evaluation models, each with his own philosophical definition. Scriven suggests evaluation is a systematic collection and analysis of information to determine the value or worth of something. Tyler focuses on the process of comparing performance data with clearly specified criteria, while Cronbach approaches evaluation from the perspective of collecting and using information to make decisions about an educational program. Gronlund views evaluation as a systematic process of determining the extent to which educational objectives are achieved by learners. Morgan and Irby’s definition refers to clinical performance evaluation as the “process of collecting information on student clinical performance in order to make informed decisions regarding student progress and program performance.”

Purpose of evaluation

Anderson and Ball state that the primary purpose of program evaluation is to determine the degree of program effectiveness. The six major purposes of program evaluation are:

1. To contribute to decisions about program installation. This includes determining if there is a need and demand for a particular program. Harless labels this as “front-end analysis,” focusing on the assessment of needs for a program, evaluation of the adequacy of the conception, estimates of cost and operational feasibility and projections of demand and support.

2. To contribute to decisions about program continuation, expansion, or certification. This purpose reflects the concept of summative evaluation or evaluation of a completed program. The intended and unintended outcomes are considered during a review of the need for a program’s continued existence.

3. To contribute to decisions about program
modification. This purpose corresponds to evaluation during the formative stages of development. The emphasis is placed on describing program processes in contrast to program products. Information derived from program objectives, content, and methodologies is reviewed and evaluated.

4. To obtain evidence favoring a program to rally support, or

5. To obtain evidence against a program to rally opposition. The realities of program evaluation are that data can be manipulated to serve motives which may be contrary to the intended purpose of a commissioned evaluation. The effectiveness of an evaluation must consider the positive and negative attributes about the program being evaluated. The presentation of evidence at its completion must be factual and unbiased.

6. To contribute to the understanding of basic processes. Two basic issues surface in addressing this purpose. First, the performance of evaluation services must be the central focus of the evaluation team. Second, limitations inherent in an evaluation must be recognized. Baker has strongly suggested that economy and parsimony be addressed in evaluation efforts. However, it is Anderson and Ball's opinion that acts of omission as well as commission are present in program planning and evaluation. Failure to address important issues sometimes leads to less than desirable results.

For the purpose of this article, attention is drawn to evaluation of student clinical performance within the framework of a nurse anesthesia program. Evaluation of clinical performance is a complex, difficult and time-consuming task which can be a source of conflict within the clinical environment. Evaluation of student clinical performance (1) verifies progress or lack of progress during the educational process; (2) improves instructor performance, due in part to student stimulation and eagerness to acquire new knowledge and skills; (3) motivates the student to improve performance; (4) assists instructors in determining the quality of student preparation for an unfamiliar learning activity; and, (5) aids in determining the adequacy of clinical sites.

Roles of evaluation

Scriven introduced the concepts of formative and summative evaluation. The terms are often erroneously classified as types or methods of evaluation. However, the terms are descriptive of the role of evaluation rather than procedures for accomplishing evaluation.

Formative evaluation relates to assessment of the quality of the instructional process or sequence and is intended to aid in the development of curriculum programs of study, teaching materials, and methods of teaching. Data gathered during a formative evaluation are aimed at improving the educational experience or product during its developmental phase. The key element in formative evaluation is feedback. Interpretation of data will often reveal deficiencies in the instructional process and allow for intervention within the learning sequence to correct the deficiencies.

Formative evaluation should be viewed as an integral component of student achievement in clinical practice. Without continual evaluation on an informal and formal basis, the student could conceivably complete an entire program without achieving the goals and objectives. Thus, the purpose is not to grade or certify, but rather, to assist the learner and instructor to focus upon the learning necessary for movement towards goal attainment.

Dick and Carey view the formative evaluation process as an opportunity to obtain data for instructors to use to increase the efficiency and effectiveness of their instructional materials. The emphasis in formative evaluation is placed on the collection of data in order to revise the instructional materials and make them as effective as possible. They have divided formative evaluation into three stages:

Stage 1: One-to-One Evaluation. Also referred to as clinical evaluation, this stage is when the instructor selects a very small group of students representative of the target population. When possible, students should all be of different general abilities. The instructional designer informs the group of the intent of the learning sequence and solicits reactions to materials, including tests. Designers of instruction have found this method invaluable. When a one-to-one evaluation is not possible, content experts can often provide the same type of feedback information needed by the designer.

Stage 2: Small Group Evaluation. After revising the materials based on the information provided in Stage 1, a group of 10-20 students is selected. Selection of students should reflect the distribution of abilities to later permit generalization of findings to the entire target population. If less than 10 students are selected, the data will not be representative of the target population. Conversely, a group of more than 20 students elicits more information than is needed. Following the
introduction and actual administration of the learning materials, the group is debriefed. Proposing several open-ended questions to the group often provides useful feedback for future revisions.

**Stage 3: Field Evaluation.** During this stage, the instructor attempts to replicate a learning situation for which the instructional materials were designed. Ideally, the group should be representative of the target population. Often times, several groups of students participate in a field-test situation to ensure that data obtained has been used under a variety of conditions.

At some point in the evaluative process, a decision must be made with regard to the effectiveness of the learning materials. This process is known as *summative evaluation*. Dick and Carey define summative evaluation as “the design, collection and interpretation of data and information.” Summative evaluation is directed toward a more general assessment in which larger outcomes have been realized over an entire course or some substantial part of a course. Thus, the main purpose of summative evaluation is to provide evidence of the effectiveness of processes and materials utilized.

The procedures followed during a summative evaluation vary depending on the purpose and focus of the evaluation. However, the techniques are similar to those used during the field-test phase of formative evaluation. It is difficult to distinguish between a formative field-test evaluation and a summative evaluation but there are two distinct differences that should be noted. The formative field-test evaluation provides data to facilitate the revision of the instructional materials, whereas following a summative evaluation, the materials usually do not undergo further modification. Summative evaluation determines the value of the presented materials for the defined group. Thus, the summative-formative distinction reflects differences in intent rather than differences in methodologies.

Formative and summative evaluation procedures are easily adapted to evaluation of student performance in a clinical situation. Student performance in a nurse anesthesia program can be divided into formative and summative evaluation time frames. A designated period can be labeled for formative evaluation, for example, mid-rotation or mid-quarter. A summative evaluation would then occur at the end of that designated time frame (at the end of rotation or of the quarter).

**Focus of clinical evaluation**

Evaluation is more than the development of effective instruments for assessment. The clinical curriculum within the framework of a nurse anesthesia program is a component of the program design. The didactic faculty emphasizes *cognitive* knowledge with consideration given to the affective and psychomotor skills. In contrast, the clinical faculty focuses on the application of the *affective* and *psychomotor* skills with the cognitive knowledge as the base.

The division between didactic and clinical evaluation is often more artificial than real in that it may be the result of physical location rather than discipline. For administrative purposes, clinical evaluation may be separated from curriculum evaluation but it should not be considered a distinct and separate entity. Clinical experience is the central activity through which the components of competence—knowledge, attitudes, skills, and judgments—are developed.

Broski developed a hierarchical model of clinical product evaluation (Figure 1). A judgment of value, worth or merit of a student’s performance is useful at any or all of the four major levels.

At the lowest level, the student receives the clinical instruction either in a classroom situation, simulation lab or as intended or unintended learning situations in the clinical setting. The student then assimilates and applies the knowledge...

**Figure 1.**

Hierarch of clinical evaluation focuses

1. **Clinical instruction**
2. **Student performance**
3. **Student product**
4. **Hypothetical of actual patient outcomes**

through performance, the second level. The third level of the hierarchy is assessment of the product or the result of the performance. The student anesthetist is evaluated on the proficiency of the given performance required during the administration of an anesthetic.

Frequently, the degree of proficiency is determined by the length of time in the program. However, the ultimate purpose of clinical evaluation is the quality of care provided to the consumer (level four). In this case, the product is viewed as a byproduct of the performance. Thus, clinical evaluation is primarily focused on performance and secondarily on the product. According to Morgan, product evaluation is "the determination of the value, merit, or worth of an artifact generated from a performance and having a functional application."

As Miller states, "The most important purpose of evaluation is really the one to which we seem to give the least conscious attention—helping students identify how much they have learned and how much they have yet to learn, in contrast to the secondary task of certifying that they have learned enough." Whether the focus of evaluation is to guide student learning or certification of competence, the beginning questions focus on what is to be evaluated. Instruction is usually designed to assist students towards a specific outcome but the instruction becomes secondary when determining if the goals have been achieved. The final test lies in what the student does, not in what was taught.

There are a variety of instruments available to assist the educator in the evaluation process. Levine indicates two processes are involved in the evaluation system. First, the data must be obtained in a manner which is both organized and quantified. Second, a decision or judgment is made based on the data. Evaluation instruments all have a number of characteristics in common. Since clinical experience focuses mainly on performance, the instruments must be designed to collect pertinent data. The choice of instruments and their method of use have a strong effect on student learning and student performance.

Positive evaluations can have significant effects on the student's self-esteem, career aspirations and career accomplishments. Clearly, a determination of what is to be measured and the type of judgment to be made as a result of the evaluation is essential. This assessment technique or combination of techniques must meet the needs of the given situation.

If measures to correct student deficiencies are not instituted, the purpose of a formative evaluation will not have been accomplished. An interaction between faculty and student should be conducted prior to the start of an educational experience for the purpose of setting goals. This allows the student and instructor to obtain a data base for the development of new skills. Evaluation should also occur at scheduled intervals during periods of learning for the purpose of assessing achievement of the stated goals and objectives. Thus, the focus of evaluation is to guide student learning and make decisions regarding student progress.

Research on clinical evaluation

Medical school students have received significant attention in the area of clinical performance evaluation. A study conducted by Benor and Hobfoll compared life experiences prior to admission to medical school to clinical performance. They found that students with recent military experience, former military officers and students with backgrounds in science performed in a way superior to younger and older students, non-officers, and non-science majors. Additionally, students who entered medical school shortly after military service performed best on clinical performance ratings. While not an unequivocal finding, former officers performed better than former enlisted men. The hypothesis that science majors did seem to have a significant edge over non-science majors was only partially confirmed in the study.

In a similar study, Murden and colleagues demonstrated that maturity, nonacademic achievement and motivation are significantly correlated with clinical performance during medical internships. An effort was made by Willoughby and associates to define complex factors related to successful clinical performance and to determine the relationship between them. A rating was completed at the close of each specified learning activity by a faculty member. The compilation of data provided the basis for judging the student's ability to "integrate and apply acquired knowledge and his attitudinal/behavioral characteristics."

The findings of this study indicate that the content areas of information, concepts and skills correlate significantly with the measures of academic achievement, whereas the content areas of attitude, peer relationships, maturity, patient rapport and integrity do not correlate significantly with the measures of academic achievement. However, the content areas reflecting personal attributes of ingenuity and conscientiousness tend to correlate significantly with the measures of academic success.
O'Donohue and Wegin studied a three-month clinical clerkship in internal medicine in which a total of 175 medical students were evaluated. Evaluations were conducted by direct observation and indicated a large intra-student variation. Individual clinical ratings had a low reliability, however, averages taken over a period of time were more consistent.\textsuperscript{21} They concluded that assessment of clinical competence is important and impacts on areas such as granting medical credentials and certification.

The results of this study did not demonstrate that written or oral examinations correlated with performance assessment; furthermore, there was considerable intra-student error within the clinical rating process. However, they did suggest reasonable judgment is essential in developing meaningful evaluation and testing techniques. Miller and associates presented evidence that showed no correlation between performance on a test of spinal anesthesia skills and a test of knowledge related to spinal anesthesia.\textsuperscript{22}

Miller has stated “that [anesthesiology] residents performed well in activities for which there were detailed learning objectives but poorly in those for which such objectives were lacking.”\textsuperscript{23} Learning objectives that include the ability to perform technical skills or develop specific attitudes are essential to develop reliable and valid methods for assessing behaviors and for training and maintaining competencies.\textsuperscript{23}

After reviewing the studies which focus on clinical evaluation, it is obvious that evaluators must consider the educational and experiential background of students in the actual evaluative process. The findings of the studies noted impact on the students' performances in the clinical environment.

**Legal implications of evaluation**

Faculty evaluation of student clinical performance is a great responsibility. Few responsibilities pose as many significant intellectual, emotional, ethical and legal demands as does the process of evaluating clinical performance. A large number of faculty members are highly reluctant to evaluate students' clinical performances with total intellectual and professional honesty. Faculty members are well aware that under the Federal Privacy Act of 1974, any student in an academic institution has the legal right to inspect and copy any records pertaining to that student which are maintained by the academic institution. In addition, faculty members are sensitive to the subjective nature of evaluating clinical performance. The intrusion of value judgments and variability among evaluators is virtually inescapable.\textsuperscript{24}

Precise methods of evaluating clinical performance varies between schools, but certain characteristics are common to all institutions. The information placed on the evaluation form by the clinical faculty member is overwhelmingly subjective, whether in the form of assigning number values to specific clinical skills or writing a narrative in response to open-ended questions.

The student may be the victim of incomplete or disingenuous evaluations. For whatever reason, evaluations often fail to inform the students about how their actual performances compare with the expected level of performance. This often occurs when there is no initiative to give the student access to the evaluation before leaving a rotation or conducting an informal conference with a faculty member.

The courts' historic respect for academic control over academic matters has been reinforced and strengthened in recent years, most notably in the U.S. Supreme Court's unanimous 1978 decision in Board of Curators of the University of Missouri v. Horowitz. Despite an exceptional academic record at the medical school, Horowitz was dismissed by the school for academic deficiency just prior to graduation. Reasons for dismissal stemmed from substandard clinical performance. In addition, it was alleged that Horowitz did not respond well to criticism. Horowitz was not given copies of formal evaluations or other evidence which the decision-making body considered during the dismissal proceedings.

The U.S. Supreme Court, in a rare 9-0 decision, rejected the student's position and approved the medical school's action.\textsuperscript{24} As indicated in the Horowitz case and other cases since, “Candor in student evaluation is permissible legally and even encouraged. Honesty is also desirable from an educational standpoint, given the uses to which clinical evaluations are put.”\textsuperscript{24}

Kapp suggests several procedural areas to ensure minimal due process and fundamental fairness for students undergoing clinical performance evaluation: (1) adopt explicit written evaluation procedures which are publicized prior to the start of the program; (2) treat students equally; (3) prepare detailed documentation of evaluations with supporting evidence for negative conclusions; (4) review evaluations prior to rotating to another area; (5) offer students the opportunity to confront the evaluator(s) informally and discuss the evaluation; (6) encourage two-way communication.
Throughout and following the rotation so that a negative evaluation is not "sprung" suddenly on a student at the conclusion of a rotation; and, (7) conduct regularly scheduled conferences with students to discuss clinical progress.24

Evaluation instruments and procedures

Data obtained in a systematic manner through the use of quality instruments and an astute faculty not only enhance the quality of information obtained but also improve the adequacy of decisions. The effectiveness of evaluation in any context depends most significantly on the instrument used in the process. Thus, the instrument becomes the focal point of communication between the student and the clinical faculty.

When the knowledge, attitudes and skills to be evaluated have been identified and the criteria for measurement established, one must consider the options available to collect data for the purpose of making a decision regarding student performance. A variety of methods are available, each having their own distinct uses, strengths and limitations. Time available to the evaluator will, in part, determine which option or combination of options is chosen.

Observation. The technique of observation is one of the most frequently used in the clinical setting. It is used to make judgments in relation to several kinds of performance—knowledge, technical skills, interpersonal skills, and habits and attitudes.25 Observation can be utilized in a variety of contexts, such as students observing other students, instructors observing students and students observing themselves. Regardless of which observation technique is adopted, recording the observations into a meaningful and useful record is essential.13

Checklist. The checklist is a systematic means of recording clinical observations. Typically, a sequential listing of phrases or objectives describe the behaviors to be evaluated. The checklist, in its true form, is not open to qualitative analysis of the selected behaviors since it demands a yes/no response. Often a rating scale is combined with the checklist to allow for judging the quality of performance.13,26

Rating scale. The rating scale is one type of format used to assess clinical behaviors and consists of two components: a set of behaviors and a graduated scale to indicate the extent to which the student is performing. The rating scale should directly relate to the learning objectives. It includes a description of the behavior and a mechanism for quantifying the behavior being evaluated.25

Anecdotal record. This is an unstructured method of recording an observation. DeMers defines an anecdotal record as "a brief description of an observed behavior that appears significant for evaluation purposes."25 The record is generally used in conjunction with another type of instrument. This format assists the evaluator in obtaining a cumulative account of supplemental information and validating student performance. Since the method is very subjective in nature, it is useful only in a formative evaluation situation.26

Critical incident technique. The critical incident technique can be used in the same manner as the checklist, rating scale or anecdotal record. The format is very similar to the anecdotal record but is more structured. Defining the behavior being evaluated becomes the incident. The analysis of the behavior is the critical aspect of this technique. It implies that the behavior has either positive or negative connotations.13,26,27

Self-evaluation. One evaluation technique which deserves additional explanation is the self-evaluation technique. This technique may be one of the most potent and useful tools any student, educator or practitioner will ever learn to use. Unfortunately, self-evaluation in the clinical environment is not used frequently nor is it used to its best advantage.

There must be an awareness of present abilities and limitations and of the areas needing further development. In many clinical practice situations, one or two nurse anesthetists provide anesthesia services for a given hospital. Because of the lack of external supervision and the trend towards increased judgment by both patients and peers, it is essential for students to learn and to develop self-evaluation skills. Unfortunately, many clinical instructors, as well as students, are the product of an educational system with a focus on external judgment rather than internal judgment. Although there is relatively limited literature available on the subject of self-evaluation, the literature demonstrates that the ability to accomplish an accurate self-evaluation is a strong determinant of psychological health. As Furhmann and Weissburg state, "Psychologically healthy persons continue to grow in their ability to evaluate themselves, make meaningful decisions and set realistic goals consistent with accurate self-concepts. The single most important determiner of self-concept is the individual's private evaluation of self,
whether that evaluation is consistent or inconsistent with external evaluations."

Individuals dependent on external feedback do not feel they have control over their environment which may cause them to mistrust their own evaluations. Subsequently, the individual becomes more dependent on this type of evaluation. And, unless individuals learn self-evaluation skills, they may not become competent, self-appraising professionals. Self-evaluation techniques can be taught and should be taught. Research indicates that the individuals who use this technique are more likely to become self-directed, self-motivated and more analytical. Self-evaluation has resulted in significantly improved work and increased critical judgment.

**Feedback.** The student needs to receive ongoing, constructive and systematic feedback in any learning situation. Many deficiencies of human performance can be related to inadequate and/or poorly managed feedback. Feedback assists the student in improving performance, acquiring additional skills and motivating the learner. A certain amount of feedback is self-administered. An inadequately managed feedback system results in a non-motivating atmosphere. According to Tucker, "Feedback acts effectively regardless of stated motivation and tends to reduce errors that can arise from a negative motivational state."

Faculty members need to recognize the role and importance of feedback. In addition to enhancing the learning environment, feedback aids in establishing an effective relationship with the learner.

Feedback should be limited to a description of the situation and relate to specific instances rather than generalizations. The focus should be on the behavior and not the personality. The time spent in a feedback situation is best if a sharing atmosphere is created. Feedback is more readily received if it is pertinent, well-timed and constructive. Imposed feedback is not as well received as when the recipient actively seeks the feedback.

This summative evaluation of the evaluation program may assist in identifying areas which need improvement, redesign and reaccomplishment. However, it is evident that an individualized program such as a nurse anesthesia program demands considerable evaluative skills on the part of faculty members and clinical instructors. Specifying objectives and measuring student performance requires continual monitoring to ascertain student progress.

**REFERENCES**


(19) Murden R, et. al. 1978. Academic and personal predic-


**AUTHOR**

Glen C. Ramsborg, CRNA, MA, is a 1965 graduate of Northwestern Hospital School of Nursing and Northwestern Hospital School of Anesthesia, Minneapolis, Minnesota. He received his Bachelor of Science in Nurse Anesthesia from the University of Minnesota in 1977 and his Master of Arts degree in Instructional Systems from the University of Minnesota in 1982. He served in the United States Air Force and was a staff anesthetist at the University of Minnesota Hospitals prior to accepting his current position as Clinical Director of the Minneapolis School of Anesthesia in St. Louis Park, Minnesota. Mr. Ramsborg is presently chairman of the AANA Continuing Education Committee and a member of the Council on Recertification of Nurse Anesthetists.

**ACKNOWLEDGEMENTS**

The author would like to thank Sandra J. Kilde, CRNA, MA, Program Director, Raymond J. Jung, CRNA, MA, Education Director and Lorraine Ferree and Ruth Radeke, secretaries, of the Minneapolis School of Anesthesia for their assistance in preparing this article.

---

**Saudi Arabia**

**professional challenges and rich rewards**

**Nurse Anesthetist**

The King Fahad Hospital in Riyadh is one of the most modern and sophisticated in the world. Having opened its first 31 beds in November 1982, it offers you the opportunity to be a member of a unique team of professionals responsible for the implementation of the facility's services. In this 500 bed acute care hospital set against the backdrop of Saudi Arabia, you can apply your skills while experiencing the richness of living in a different and dynamic culture.

Reporting to the Anesthesiologist, your duties will include the administration of intravenous, spinal and other anesthetics during surgery, deliveries and other medical and dental procedures.

... then meet the challenges and experience the rewards by sending your resume, in confidence, to:

- **Heather A. Campbell,**
  Jonathan Walsh & Associates, Ltd.,
  Ref. 1413/AANA/02,
  Two Robert Speck Parkway,
  Suite 1410,
  Mississauga, Ontario L4Z 1H8, Canada

---

**Saudi Arabia**

**professional challenges and rich rewards**

**Nurse Anesthetist**

You must be a graduate of an accredited school of nursing and an accredited school of nurse anesthesiology, certified as a Nurse Anesthetist by the AANA and have at least two years experience as CRNA. RN and Nurse Anesthetist licensure is also required.

**If you are...**

Eligible to travel on a U.S. Passport

Prepared to accept a two year contract with furnished married status accommodation and a full range of benefits

---

**JONATHAN WALSH & ASSOCIATES LTD**

**QUALITY CARE**

Authorized Representative

HCA International Ltd.

---

**Journal of the American Association of Nurse Anesthetists**