*Athletic Training Education Journal 2009;4(3):109-116*

www.nataej.org

# Issues in Selecting Methods of Evaluating Clinical Competence in the Health Professions: Implications for Athletic Training Education

David A. Middlemas, EdD, ATC*; Carleton Hensal, ATC†
*Montclair State University, Montclair, NJ; †High Point Regional High School, Sussex, NJ

**Objectives:** To examine methods used to evaluate the clinical competence and proficiency of students in medicine and allied health professions. To identify factors that would be valuable to educators in athletic training and other medical and allied health professions in the development and use of clinical assessment methods.

**Data Sources:** We searched EBSCO Academic, InfoTrac, MEDLINE and PubMed for literature from professional journals related to clinical education, assessing clinical competence, clinical competence, clinical proficiency assessment.

**Data Synthesis:** Assessment of competence in the clinical setting is an issue for medicine and allied health education programs, including athletic training. Methods used to assess clinical competence included written, objective tests, checklists, oral examinations, patient management problems, simulated patients, observed clinical situations, the Observed Structured Clinical Exam and observation of students. Each method has its advantages and concerns. Psychometric issues relating to reliability, validity and generalizability arise in the development and implementation of clinical assessment instruments. In addition, there are concerns about how faculty expectations relate to student performance and how novices perform in the clinical setting as compared to experienced practitioners.

**Conclusions/Recommendations:** Although specific outcomes may differ between the medical and allied health professions, athletic training educators can benefit from the lessons learned about methods used to evaluate students in the clinical setting. In an effort to capitalize on the benefits and minimize the problems with the assessment of clinical competence and proficiency, athletic training educators must be aware of the psychometric and outcome-related issues in test development. Attention should be paid to issues of reliability, validity and generalizability when developing tests. Clinical education outcomes must be developed in a manner that addresses the variety of cognitive levels at which clinicians function in professional practice to develop authentic assessments of clinical proficiency.

**Key Words:** clinical education, clinical competence, athletic training education, assessment, patient simulation

Changes in athletic training education and revision of the criteria for accrediting athletic training education programs (ATEPs) have encouraged faculty to develop outcomes for clinical education that are competency based. The changes and revisions have placed emphasis on the assessment of the clinical competence and proficiency of athletic training students in a framework that demonstrates students are "learning over time."[1-4] To this end, it would seem appropriate to identify and implement valid and reliable methods of assessing outcomes that indicate the students' competency and proficiency in the clinical setting.

*Dr. Middlemas is an Associate Professor and the Athletic Training Program Director at Montclair State University*
*middlemasd@mail.montclair.edu*

*Mr. Hensal is an athletic trainer at High Point Regional High School.*

The National Athletic Trainers' Association Education Council (NATAEC) states that "These competencies provide the entry-level ATC® with the essential knowledge and skills needed to provide athletic training services to patients of differing ages and genders and work, and lifestyle circumstances and needs."[5] The list of behavioral objectives or outcomes, is categorized into the following behavioral classifications: Cognitive Competencies, Psychomotor Competencies and Clinical Proficiencies.[5]

The NATAEC has stated that the entry-level clinical proficiencies "serve two purposes: (1) they define a common set of skills that entry-level athletic trainer should possess; and (2) they redefine the structure of clinical education from the quantitative approach (i.e., "hours") to an outcomes-based qualitative system."[6 p.iv] This suggests the need for students to integrate information and clinical skills from the twelve content areas of the athletic training profession into the clinical decision making process. As a result, ATEPs should assess at the highest levels of the cognitive and psychomotor taxonomies.

The use of outcome-based educational models is justified since the ability to predict performance on the Board of Certification (BOC) examination from the amount of time spent in the clinical setting is not supported in the literature.[7,8] As allied health profession faculty, athletic training educators would be aided by an understanding of the lessons learned about the various methods of assessing the clinical competence of students in medicine and allied health that have already been discussed in the literature.

Information in the athletic training literature relating to clinical education deals with learning styles,[9-11] critical thinking[12,13] guidelines for clinical education settings,[14] student perceptions of characteristics of clinical instructors,[15,16] assessment of clinical instructor behavior[17,18] and clinical education models.[19] The athletic training education literature has presented methods of assessing clinical competency at varied levels of complexity[20-23] ranging from checklists[21] to visual scales, to open-ended questions, to proficiency rating scales on which the clinical instructor provides feedback.[20-23]

Athletic training educators continue to search for ways to evaluate the clinical competency and proficiency of their students within the framework of outcome-based clinical education. In 1997 Starkey[24 p. 114] stated, "Our students' clinical education model should be based on a set of measurable, standardized, and referenced learning objectives that describe the type and nature of the experience obtained." The logical next step would be to assess students with valid and reliable methods that provide feedback on the learning taking place during that experience.

As a potential first step in facilitating the process of developing appropriate tools for the assessment of clinical proficiency, we felt it valuable to investigate the methods used in medicine and other allied health professions to assess student performance during the clinical portion of their education. Our purpose in doing so was to identify common strengths, weaknesses, and issues related to the process of clinical competency assessment in medicine and allied health professions.

## Assessing Clinical Competence

As early as 1992 the literature identifies issues relating to the assessment of professional preparation of athletic trainers in the clinical setting. Weidner and Vincent[24] reported that only about 20% of newly employed athletic trainers strongly agreed that their clinical instructors provided them with a realistic impression of their overall professional readiness to enter the job market. There were similar attitudes reported about the nature of "mentoring". Recommendations at that time included improving and increasing the mentor relationship with more emphasis on professional evaluation of the students.[25]

Issues and questions concerning assessment of students in the clinical setting are not unique to athletic training education. Faculties in medicine, dentistry, nursing and physical therapy have been wrestling with this problem for years.[26-34] In medicine patient management problems (PMPs) were developed during the 1960s to attempt realistic evaluation of medical problem solving. Medical licensing examinations discontinued these assessment instruments, which used latent-image pens, in the 1980s when psychometric problems were identified.[26]

One intent of performance-based assessment is to measure one or more aspects of higher order cognitive processes. Student responses to surface features of performance may not guarantee that higher order thinking processes occur.[29] In addition, instructional experiences can undermine the intent to measure higher order thinking.[29-30] Relating this to the assessment of clinical competence and proficiency in athletic training, the educator should be aware that instruments designed to measure clinical competence may not, in fact, measure the construct of "competence," but some lower order thinking process such as knowledge of facts or procedures.[28] When one considers that performance assessment is complicated logistically and technically, serious problems, including potential corruption of the process, can arise particularly if the assessment involves a high stakes situation.[27-28]

Athletic training students who become candidates for BOC certification must demonstrate proficiency on educational competencies in the cognitive and psychomotor domains.[5] Kolb and Shugart[31] stated that most of the evaluation instruments in current use measure cognitive and psychomotor skills with few instruments measuring affective behavior. They also stated that one can find only an occasional tool that attempts to measure skills in all three behavioral domains at the same time.

Assessing the clinical competence of medical and allied health students is an issue facing educational programs throughout health care. Although specific outcomes may differ among disciplines, faculty in medicine and the allied health professions have used a variety of assessment methods in an effort to determine clinical competency of their students. Each of the methods has its advantages and problems when used in the clinical education setting. We present and discuss a summary of these issues below.

### Multiple Choice Questions and Objective Tests

Medicine has used multiple choice questions (MCQs) in the clinical education settings. One study reported that MCQs were one of the three most commonly used methods of assessment of factual knowledge and abstract problem solving.[36] It has been reported that MCQs only measure knowledge of facts and information and the test takers' ability to recite or recognize information.[33] The levels of cognitive processing discussed by Van der Vleuten and Newble[33] are consistent with the lowest two levels of Bloom's taxonomy, "knowledge" and "understanding", or the lowest levels of Miller's Pyramid of Clinical Competence,[35] "knows" and "knows how."

Evaluation of factual knowledge and problem-solving skills with MCQs appear to offer excellent reliability. Scores on MCQ exams have also demonstrated a high correlation with subsequent

faculty and peer ratings,[37,38] but their validity has been questioned.[34,37] The high correlation has been used to hypothesize that the MCQ may assess some aspect of context and clinical reasoning.

In a study comparing medical students in traditional medical schools with students in innovative programs on whether written assessments of skills could predict performance on the objective structured clinical examination (OSCE), significant differences with the performance on MCQs were found between students attending innovative medical schools and those in more traditional programs.[39] The authors argued that a specific trait, such as the knowledge of clinical skills and not general medical knowledge, was measured using written tests. These results question the construct validity of written skills tests.[33,39]

The ability of MCQs to measure the knowledge of clinical skills may provide feedback on some aspect of clinical reasoning. Without evidence to support a strong relationship between the performance on MCQs and clinical skill performance, there remain problems with their use in the clinical setting as they do not measure student performance in the psychomotor domain.

**Checklists**

Initially in athletic training education, assessment models that used checklists for the evaluation of student performance were designed to meet the accreditation requirements of the Joint Review Committee of educational programs in Athletic Training (JRC-AT); then under the egis of the Commission on Accreditation of Allied Health Education Programs (CAAHEP). The use of skill oriented checklists was similar to the format of the practical portion of the athletic training certification examination in use at the time.[40-41]

There are inherent problems with the use of checklists for the assessment of clinical competence.[20,22] Student scores on checklists may tend to increase with increased expertise of the examiner.[42] In addition, student scores on checklists may also be affected by an evaluator's perception of the reputation of an education program.[43,44] It is also possible that external raters on checklists may not be attuned to the affective components of a clinical competency or its performance.[34,42, 43]

Student scores on checklist-based instruments may also be affected by conversations between evaluators after completion of an examination,[45] a factor identified in the BOC's instructions to examiners when the practical examination was in use. Checking boxes and marking scales may not contain details of an observed performance and provide no opportunity for the observer or faculty member to make judgments related to clinical performance.[38,41]

Although checklists may be appropriate for recording whether a student has completed assigned skills, their use for assessing the construct of clinical competence may not provide the type of feedback sought by the evaluator. The concerns related to reliability and examiner bias may skew the results of the evaluation enough to affect its usefulness.

**Oral Examinations**

Methods for oral examinations include station examinations, examiner role playing and examinees taking oral questions from patient charts.[47] Exams can last from 30 minutes to two hours with as many as one to five examiners.[27]

Medical licensing examiners eliminated oral examinations from their tests in the early 1960s for logistical and psychometric reasons after attempts to increase their reliability failed. Analysis of performance indicated that there were near-chance levels of agreement between the examiners.[27] It has been reported that oral examinations lack the range of cases and open ended types of problem resolutions to truly assess clinical competence.[46]

Swanson[26] reported that many medical schools still used oral examinations with modifications in order to standardize the examinations as much as possible. In contrast, Mavis[48] reported in a study of 126 allopathic medical schools that oral examinations were not used to determine grades for the majority of courses or to help determine graduation. The study also stated that only 21% of the schools indicated using oral examinations to determine clerkship grades.

The generalizability of oral examination scores tends to be low. Generalizability coefficients between .22 and .87 have been reported.[49] Improvement has been seen with increases in testing time or exposures. One serious drawback may be that the increases in testing time to longer-than-a-day were required to bring the coefficients to acceptable levels.[26,32,49]

Reliability of unstructured oral exams is extremely low.[50] Some of these potential reasons include the subjectivity of the examiner, a small number of cases being presented to the candidate, and personal feelings between the examiner and subject. In addition, the possible existence of the halo effect (i.e., the tendency to extend the impression of positive performance on one behavior to the evaluation of other behaviors or the overall assessment of the individual) can affect scores. Proposed solutions to reliability problems with oral examinations include testing longer and increasing the sample.[49,51] It is possible that both these methods will not solve the problem because increasing test length will not necessarily cause the halo effect to disappear. Also, the lengthening of tests may extend into multiple days. In the view of the teacher, the amount of time required to do this may make these changes unfeasable.[51]

**Patient Management Problems**

The patient management problem is a sequential, problem-based testing method that uses a written instrument to disclose data to the examinee through technical means, such as the latent image pen. The examinee's score is based on the decisions he/she made and the pathway he/she follows to these decisions.[33] The PMP was developed in the 1960s to pose more realistic challenges to problem solving skills, largely in response to dissatisfaction with traditional, factually oriented, MCQ tests. They tend to involve an opening scenario and a series of scenes through which the examinee proceeds, gathering information and initiating patient management activities. The examinee selects

clinical actions to which he/she receives feedback. Typically the problems take 10 to 30 minutes to complete depending on their complexity.[27]

Issues with PMP include difficulty in obtaining consensus on the positive and negative weights assigned to interventions and the optimal path to the solution of the problem. In addition, the exams are confounded by the virtually unavoidable element of cueing and unintentionally providing the examinee with information which helps to solve the problem when the exams are in a written form.[36] They were discontinued from medical licensing exams in the 1980s due to widespread psychometric problems.[25,33,50]

The BOC written simulation examination used a latent image response form of the PMP to test candidates. The BOC reported that the examination was content valid based on the Role Delineation Study, fourth edition.[52] In 1999 the BOC reported reliability measures of internal consistency using the Kuder-Richardson Formula 20 of .89.[52] In 2002 measures of internal consistency for the written simulation were reported as .95 and .96 for the two forms of the examination.[53]

Again, the question arises as to whether or not a written assessment method of a psychomotor component is a valid measure of the construct of clinical competence or proficiency. There are no reports in the literature regarding the assessment of the clinical or professional competence construct validity of the examination scores for the BOC written simulation examination.

**Simulated Patients**

The use of simulated patients (SPs) for the assessment of student competence in clinical education is well debated in medical and allied health education.[27,31,44,54-60] The use of SPs appears to have begun in the 1960s and 1970s when clinical reasoning was thought to be a construct. Educators accepted using SPs for student assessments more due to the perception that there was an increase in realism than due to its validity and reliability.[33]

This method of assessing students involves the use of trained SPs to assess proficiency in the clinical setting. In 1987 Barrows[54] defined the simulated patient as a person who has been carefully coached to accurately portray a specific patient when given the details of the history and physical examination. The students' performance is evaluated either by the SP or by one or more faculty members. Information supporting the use of SPs and its value is balanced with a discussion of problems and/or concerns including issues relating to the reliability of assessment scores.

Proponents of using assessment with SPs have argued that there is greater realism provided in clinical simulations than in paper-and-pencil exams, especially in the affective and psychomotor domains.[55,58,61] In addition, the possibility exists that students can be evaluated on multiple components of clinical proficiency, such as observational skills, didactic knowledge, problem solving, clinical and technical skills, decision making and interacting with patients.[31,54,55] The tendency of this assessment method being realistic, having good construct validity and having acceptable reliability are also factors that support the use of SPs as evaluation tools.[31, 54] Evaluation with SPs appears to be effective for assessing competencies that require direct interaction with the patient.

The ability of the evaluator to control the variables of the testing situation, hence eliminating some of the uncertainty in the clinical setting, is an advantage of using SPs. The logistics of the assessment method also appear to be manageable for interpersonal skills like history taking, communication and direct interaction with the patient.[31,55]

Concerns about the assessment of students using clinical simulations were noted as early as the 1960s and 1970s.[58] Problems identified during this time were related to clinical problem solutions, data collection processes, time allocation decisions, resources and reliability. Questions related to item weighting and scoring have also been discussed.[34,58,61]

Student assessments using SPs place a high demand on available time and resources, with potentially negative impacts on the instruments and their administration.[27,55,56,60,61] The high demand on resources and time is compounded by the cost of training simulated patients for assessments.[27,56,61] It has also been stated that the exams lack the appropriate range of cases for the proper assessment of students.[31,32,46,49,56,57]

Score reliability analysis has supported the reproducibility of student performance on individual cases. Although reported intrarater reliability coefficients of .77-.91 tended to be lower than interrater coefficients of .82-.90,[60,62] they were high enough to argue that students performed consistently on the same cases. On the other hand, concerns about reliability have arisen because there is some evidence that performance of a student on one case is not very predictive of his/her performance on another case.[56] Low correlation coefficients between student scores on different cases refutes the argument that clinical reasoning is a generic ability.[33,46,59,60,62]

Standardization of the assessment process with SPs is an issue in potential need of attention. Introduction of factors like cueing, the conscious or unconscious giving of feedback during the process, presents a problem because the test may not discriminate among the students any better than the MCQ. Also, the uncontrolled nature of the clinical environment can affect the standardization of the testing situation with respect to the patients and the observers. These factors can contribute to low test reliability.[33,55,59,60] Although standardized training for the SP and raters may help increase reproducible and consistent scenarios, the benefits must be weighed against the issues related to allocation of resources.

**Observed Clinical Situations**

*In vivo* assessment of students in the clinical setting tends to be unstructured and not standardized. Issues of reproducibility and standardization arise because encounters with patients are insufficiently reproducible or predictable. In the clinical setting, real patients may not necessarily match the level of knowledge

evaluated in the assessment. It is also possible that reliability is affected by the inability to separate student performance from other variables in the clinical situation.[49,57] The literature also identifies issues such as the halo effect or professional judgment as factors affecting the nature of the students' scores on assessment instruments.[47,57]

The amount of time involved for faculty to complete *in vivo* observations is an issue. The observation of complex clinical situations demands long and impractical exams that can possibly last longer than a day. The amount of faculty enthusiasm for doing observations can affect student scores on observation assessments.[34,57]

**The Observed Structured Clinical Exam and Observation of Students**

The OSCE is a timed multi-station examination often using SPs to simulate clinical scenarios. In the OSCE the roles are portrayed accurately and convincingly.[39] It is a specific example of the SP examination used in medicine. Nursing and physical therapy also use the OSCE. The exams concentrate on hands-on clinical behavior where assessors ask students to show how to perform clinical tasks.[59,63]

There is mixed feedback on the OSCE. The literature reports a high correlation between written and clinical examination formats, which challenges the concept that the format dictates what is being measured.[34] Scores on written tests can predict student performance on the OSCE. Although there appears to be a relationship between scores on written measures and the OSCE, its apparent face validity may make it useful when assessing the effectiveness of medical schools.[40]

There is some support for the use of the OSCE in the medical literature. It has been hypothesized that schools using the OSCE as part of their curriculum would have a particular advantage over students from schools that do not use it.[40] The examination concentrates on hands-on clinical behavior in which examiners ask students to "show how" to perform the select behavior. However, as originally proposed, the exam measures clinical skills in isolation and over short periods, which is not a valid representation of clinical reality at the end of a curriculum.[62] There is also support for the need to demonstrate of skills in isolation as the knowledge of skills is important for the proper performance of skills.[52]

A three year experience with the pediatric OSCE exam found it to be feasible, albeit labor intensive and costly. Joorabchi and DeVries[63] felt it could be implemented with high reliability and validity. They also noted a wide gap between faculty expectations and student performance.

**Patient Simulators**

Some of the latest technology in the instruction and evaluation of clinical skills and decision making involves the use of human patient simulators (HPSs). The simulators are fully computerized mannequins programmed to teach pathophysiology, pharmacology and comprehensive scenario testing of one or more critical health incidents.[64] Documentation of the use of HPS mannequins goes back to 1960 with the introduction of Resusci Annie® for the instruction of cardiopulmonary resuscitation (CPR). Development of the Recording Annie in 1971 provided one of the first mannequins that gave instant feedback about the quality of the skill performance.[65] Today HPSs are used to instruct and assess skills including CPR, defibrillation, emergency cardiac care, endotracheal intubation, patient assessment and surgery.

The literature supports the use of simulators for training.[66-70] Along with the ability to practice and evaluate skills and scenarios without putting patients at risk, evaluations of interns who were trained on HPSs indicated significantly higher performance on skills in real emergency situations than those who were not trained on HPSs.[66-68,70] Team members trained on simulators were reported to have higher confidence in performing skills[68] and better adherence to American Heart Association standards for cardiac arrest.[69]

The use of HPSs for training and evaluation of patient care skills also raises some questions. In a study evaluating the effectiveness of an HPS for teaching the evaluation of ventricular heart disease, there was no significant difference in the performance of the control group to the experimental group, who were trained on the simulator.[71] The author stated that the intensity or duration of the training may have contributed to the result. In a study of the effectiveness of computer and simulator training for CPR, the value of computerized and simulator training in lieu of traditional classroom instruction for initial training may be questionable as skill performance for the group trained by the computer and simulator was worse. The use of computer and simulator may be more effective for recertification.[72] These authors both indicated a need for further research into simulator-based instruction and evaluation of skills.

Additional time demands for faculty are also a factor when considering the implementation of training and evaluation with the HPS. It has been reported that a large amount of time is required for the development of scenarios for training and evaluating students.[73] A significant factor that would affect the implementation of HPS into athletic training education programs is cost. Prices range from $3800.00 for relatively basic simulators[74] to over $150,000 for more complete computerized HPS.[75] The possibility of pooling resources or sharing simulator laboratories might make the technology available for programs that could otherwise not afford them.

## Conclusions

Implementing effective methods of evaluating clinical competency and proficiency is an issue that is common to education programs in athletic training, medicine and other medically related professions. Although specific clinical outcomes may vary among the professions, the process of evaluating student performance on those outcomes has resulted in questions and concerns that are somewhat consistent across

the disciplines. When developing and revising clinical assessment tools, athletic training educators can benefit from the lessons learned by those in the various disciplines. There are many methods used to assess the performance of students in the clinical setting. While each may have its advantages, there are also limitations that may affect using the results of any specific assessment method as the sole determinant of a student's clinical proficiency. It is important for the educator to weigh the positives and negatives of the various methods in relation to the outcomes assessed before selecting the method that will best meet the needs of a particular situation.

Examples of positive factors identified in the literature include the use of multiple choice questions to evaluate a student's specific knowledge or understanding of a clinical concept, checklists to record student completion of specific skills and procedures, and methods employing simulated patients or observed clinical situations to evaluate the student's ability to assesses, diagnose and provide appropriate patient care. Some of the latest technological advances in patient simulators have the potential to provide educators with the ability to obtain objective data when evaluating higher risk skills and proficiencies without risking patient safety.

Patient management problems, simulated patients, observed clinical situations and the OSCE are examples of methods developed to address issues related to evaluating clinical decision making and more complex clinical skills. The thought is that these methods are better suited to assess critical thinking, clinical decision making, cognitive processing and problem solving. All of these factors are potential components of the construct of "clinical proficiency" in athletic training.

In working to develop assessment methods that appropriately evaluate clinical proficiency outcomes, educators should be aware of the common issues with these methods. An excellent starting place is to identify and address psychometric factors such as validity, reliability and generalizability, which are concerns for the majority of the clinical education assessment methods.

In addition, administration methodology and security have affected the design and implementation of evaluation methods. The time and cost of training evaluators and models, as well as the test itself, also present formidable obstacles for appropriately investigating and using some methods. Finding ways to successfully address these issues, especially those related to time and cost, can help educators and researchers investigate and refine assessment tools to better evaluate clinical competence and proficiency.

Miller's pyramid of clinical competence illustrates one theory on the progression of clinical proficiency. At the lowest level of the pyramid is knowledge (knows), followed by competence (knows how), performance (shows how), and action (does).[72] Reliable and valid evaluations of the "knows" and "knows how" levels can provide information about the students' ability to access information and demonstrate skills in the practice of athletic training. Assessment of the "shows how" and "does" levels may provide a more valid and comprehensive picture of clinical competence and/or proficiency. Case-based evaluation of student performance attempts to target these higher levels of the pyramid, potentially providing insight into performance of the individual in professional practice.

The responsibility of the ATEP to verify student clinical proficiency has bolstered the need to continue investigating assessment models for clinical education. The elimination of the practical portion of the certification examination and requirements for programs to verify the assessment of clinical competencies and proficiencies are strong motivators for athletic training educators to design, revise, validate and implement methods of student assessment that provide the most accurate and useful information.

## Recommendations

Athletic training education could be enhanced by developing psychometrically sound formative and summative assessment models for clinical education. Development and investigation of outcome-based assessment models for clinical education that address cognitive and psychomotor taxonomies, critical thinking, multiple intelligence and clinical proficiency may set a standard for the assessment of clinical competence in multiple disciplines.

Identification and verification of the factors that make up the construct of clinical proficiency will help to better define the foundation upon which clinical assessment models are designed. In addition, a better understanding of the construct will help educators insure that the educational outcomes currently in place are accurate indicators of appropriate professional education.

## References


1. Commission on Accreditation of Allied Health Education Programs. *Standards and Guidelines for the Athletic Trainer*. Chicago, IL: Commission on Accreditation of Allied Health Education Programs; 2001: 8-10.
2. National Athletic Trainers' Association. Learning over time models. In: *Clinical Instructor Educator Seminar Workbook*. Dallas, TX: National Athletic Trainers' Association; 2001: viii,8-9.
3. National Athletic Trainers' Association. Teaching and evaluating the clinical proficiencies. *NATA News*. July 2002:21.
4. Joint Review Committee on Educational Programs in Athletic Training. *Interpretation Manual for the 2001 Accreditation Standards and Guidelines.* Centennial, CO: Joint Review Committee on Educational Programs in Athletic Training; 2001:37,41,50-52.
5. National Athletic Trainers' Association. *Athletic Training Educational Competencies.* 4th ed. Dallas, TX: National Athletic Trainers' Association, Inc.; 2006.
6. National Athletic Trainers' Association. *Athletic Training Educational Competencies.* 3rd ed. Dallas, TX: National Athletic Trainers' Association, Inc.; 1999.
7. Middlemas DA, Manning JM, Gazzillo LM, Young J. Predicting performance on the National Athletic Trainers' Association Board of Certification examination from grade point average and number of clinical hours. *J Athl Train.* 2001;36:136-140.
8. Harrelson GL, Gallaspy JB, Knight HV, Leaver-Dunn D. Predictors of success on the NATABOC certification examination. *J Athl Train*. 1997;32:323-327.

9. Draper DO. Students' learning styles compared with their performance on the NATA certification exam. *Athl Train*. 1989;24:234-235,275.
10. Harrelson GL, Leaver-Dunn D, Wright KE. An assessment of learning styles among undergraduate athletic training students. *J Athl Train*. 1998;33:50-53.
11. Coker CA. Consistency of learning styles of undergraduate athletic training students in the traditional classroom versus the clinical setting. *J Athl Train*. 2000;35:441-444.
12. Fuller D. Critical thinking in undergraduate athletic training education. *J Athl Train.* 1997;32:242-247.
13. Leaver-Dunn D, Harrelson GL, Martin M, Wyatt T. Critical-thinking predisposition among undergraduate athletic training students. *J Athl Train*. 2002;37(suppl 4): S-147–S-151.
14. Weidner TG, Laurent T. Selection and evaluation guidelines for clinical education settings in athletic training. *J Athl Train.* 2001;36:62-67.
15. Laurent T ,Weidner TG. Clinical instructors' and student athletic trainers' perceptions of helpful clinical instructor characteristics. *J Athl Train.* 2001;36:58-61.
16. Stemmans CL, Gangstead SK. Athletic training students initiate behaviors less frequently when supervised by novice clinical instructors. *J Athl Train*. 2002;37(suppl 4):S-255-S-2260.
17. Weidner TG, Henning JM. Importance and applicability of approved clinical instructor standards and criteria to certified athletic trainers in different clinical education settings. *J Athl Train.* 2005;40(4):326-332.
18. Stemmans CL, Gangstead SK. A psychometric investigation of the clinical instruction analysis tool-athletic training II. *Percept Mot Skills.* 2000;91:1040-1044.
19. Radtke S. A conceptual framework for clinical education in athletic training, *Athletic Training Education Journal*. 2008;3(2):36-42.
20. Evaluation of the athletic training clinical proficiencies: a modular approach. As found at http://www.nata.org/student/clinicalmod.htm. Accessed July 9, 2002.
21. Amato HK, Konin JG, Brader H. A model for learning over time: the big picture. *J Athl Train*. 2002;37(suppl 4):S-236-S-240.
22. Amato HK, Venable Hawkins C, Cole SL. *Practical Exam Preparation Guide of Clinical Skills for Athletic Training*. Thorofare, NJ: Slack; 2002.
23. Knight, K. *Assessing Clinical Proficiencies in Athletic Training: A Modular Approach*. 3rd ed. Champaign, IL: Human Kinetics; 2001.
24. Starkey, C. Reforming athletic training education. *J Athl Train*. 1997;32:113-114.
25. Weidner TG, Vincent WJ. Evaluation of professional preparation in athletic training by employed, entry-level athletic trainers. *J Athl Train.* 1992; 27:304-310.
26. Swanson DB, Norman GR, Linn RL. Performance-based assessment: lessons from the health professions. *Educ Researcher*. 1995;24(6):5-11,35.
27. Brandt R. On performance assessment: a conversation with Grant Wiggins. *Educ Leadership*. 1992;49(8):35-38.
28. Baker EL, O'Neil HF, Linn RL. Policy and validity prospects for performance-based assessment. *Am Psychol*. 1993;48(12):1210-1218.
29. Wiggins G. A true test: toward more authentic and equitable assessment. *Phi Delta Kappan*. 1989;May:703-713.
30. Cross V. Approaching consensus in clinical competence assessment. *Physiotherapy*. 2001;87(7);341-350.
31. Kolb SE, Shugart EB. Evaluation: is simulation the answer? *J Nurs Educ.* 1984;23(2):84-86.
32. Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. *Lancet*. 2001;357:945-949.
33. Van der Vleuten C, Newble D. How can we test clinical reasoning? *Lancet.* 1995;345:1032-1034.
34. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA.* 2002;287(2):226-235.
35. Miller GE. The assessment of clinical skills/ competence/ performance. *Acad Med*. 1990;65:563-567.
36. Newble D, Dawson B, Dauphinee D. Guidelines for assessing clinical competence. *Teaching Learning Med.* 1994;6:213-220.
37. Donwing SM. The assessment of clinical competence on the emergency medicine specialty examination: the validity of clinically relevant multiple-choice items. *Ann Emerg Med.* 1980;9:554-556.
38. Ramsey PG, Wenrich MD, Carline JD, Inui TS, Larson EB, LoGerfo JP. Use of peer ratings to evaluate physician performance. *JAMA.* 1993;269:1655-1660.
39. Remmen R, Scherpbier A, Van der Vleuten C, et al. Effectiveness of basic clinical skills training programmes: a cross-sectional comparison of four medical schools. *Med Educ.* 2001;35:121-128.
40. NATABOC frequently asked questions: exam development questions. National Athletic Trainers' Association Board of Certification, Inc. http://www.nataboc.org/atc/links/faq/. Accessed March 5, 2003.
41. Columbia Assessment Services. *Facilitator Training Manual for the Examiner Training Program*. Omaha, NE: National Athletic Trainers' Association Board of Certification, Inc.; 2000.
42. Blane CE, Calhoun JG. Objectively evaluating student case presentations. *Invest Radiol*. 1985;20:121-123.
43. Van der Vleuten C, Swanson DB. Assessment of clinical skills with standardized patients: state of the art. *Teaching Learning Med*. 1990;2:58-76.
44. Stillman P, Swanson D, Regan MB, et al. Assessment of clinical skills of residents utilizing standardized patients: a follow-up study and recommendations for application. *Ann Intern Med*. 1991;114:393-401.
45. Newble DI, Swanson DB. Psychometric characteristics of the objective structured clinical examination. *Med Educ*. 1988;22:325-334.
46. Cox K. Examining and recording clinical performance: a critique and some recommendations. *Educ for Health: Change in Learning and Practice*. 2000;(13):45-53.
47. Stratford PW, Thomson, MA, Sanford J, et al. Effect of station examination item sampling on generalizability of student performance. *Phys Ther*. 1990;70(1):45-50.
48. Mavis BE, Cole BL, Hoppe RB. A survey of information sources used for progress decisions about medical students. *Med Educ Online.* 2002;7:12. http://www.med-ed-online.org. Accessed July 8, 2007.
49. Daelmans HEM, Scherpbier AJJA, Van der Vleuten CPM, Donker ABJM. Reliability of clinical oral examinations re-examined. *Med Teacher*. 2001;23(4):422-424.
50. Swanson, DB, Norcini JJ, Grasso L. Assessment of clinical competence: written and computer-based simulations. *Assessment and Evaluation in Higher Education*. 1987;12:220-246.
51. Schwirth LWT, Van der Vleuten CPM. Quality control: assessment and examinations. *Quality of Teaching in Higher Education*. 1996;1(2):66-76. http://www.oeghd.or.at/zeitschrift/1996h1-2/06_art.html. Accessed September 24, 2007.
52. National Athletic Trainers' Association Board of Certification. *Role Delineation Study*. 4th ed. Omaha, NE: National Athletic Trainers' Association Board of Certification, Inc.; 1999.
53. Castle Worldwide. *Procedures Used to Develop the National Athletic Trainers' Association Board of Certification Examinations*. Morrisville, NC: Castle Worldwide, Inc.; 1999. http://www.nataboc.org. Accessed March 1, 2003.
54. Barrows H. Simulated (standardized) patients and other human simulations. Presented at: Health Sciences Consortium; 1987; Chapel Hill, NC.

55. Ladyshewsky R, Baker R, Jones M, Nelson L. Evaluating clinical performance in physical therapy with simulated patients. *J Phys Ther Educ*. 2000;14: 31-37.
56. Stillman PL, Regen MB, Swanson DB. A diagnostic fourth-year performance assessment. *Arch Intern Med*. 1987;4:1981-1985.
57. McDowell BJ, Nardini DL, Negley SA, White JE. Evaluating clinical performance using simulated patients. *J Nursing Educ.* 1984;23(1):37-39
58. McGuire CH, Babbott D. Simulation technique in the measurement of problem-solving skills. *J of Educ Measurement*. 1967;4(1):1-10.
59. Stillman PL, Swanson DB. Ensuring the clinical competence of medical school graduates through standardized patients. *Arch Intern Med*. 1987;147:1049-1052.
60. Ladyshewsky R, Baker R, Jones M, Nelson L. Reliability and validity of an extended simulated patient case: a tool for evaluation and research in physiotherapy. *Physiotherapy Theory and Practice.* 2000;16:15-25.
61. Vu NV, Barrows HS. Use of standardized patients in clinical assessments: recent developments and measurement findings. *Educ Researcher*. 1994;April:23-30.
62. Van der Vleuten, C. Validity of final examinations in undergraduate medical training. *BMJ*. 2000;321:1217-1219.
63. Joorabchi B, Devries JM. Evaluation of clinical competence: the gap between expectation and performance. *Pediatrics.* 1996;97:179-184.
64. Nehring WM, Lashley FR. Current use and opinions regarding human patient simulators in nursing education: an international survey. *Nursing Education Perspectives.* 2004;25(5):244-248.
65. Tjomsland N, Baskett P, Asmund S. *Laerdal Resuscitation*. 2002; 53(2):115-119.
66. DeVita M A, Schaefer J, Lutz J, Wang H, Dongilli T. Improving medical emergency team (MET) performance using a novel curriculum and a computerized human patient simulator. *Qual Saf Health Care*. 2005;14(5):326-31.
67. Park J, MacRae H, Musselman LJ, Rossos P, Hamstra SJ, Wolman S, Reznick RK. Randomized controlled trial of virtual reality simulator training: transfer to live patients. Am J Surg. 2007;194(2):205-211.
68. Seybert AL, Barton CM. Simulation-based learning to teach blood pressure assessment to doctor of pharmacy students. *Am J Pharm Educ*. 2007;71(3):48.
69. Wayne DB, Didwania A, Feinglass J, Fudala MJ, Barsuk JH, McGaghie WC. Simulation-based education improves quality of care during cardiac arrest team responses at an academic teaching hospital: a case-control study. *Chest*. 2007; 1: 7-131
70. Adachi M, Ribaudo V, Rosenthal M, Tsegaye A, Schneider R, Mayo PH. In-hospital cardiopulmonary resuscitation (IHCPR): results of a program to improve team leader skills using a computerized patient simulator (CPS). *Chest* 2005;128(4):182-183
71. Oddone EZ, Waugh RA, Corey R, Feussner JR, Samsa G. Teaching cardiovascular examination skills: results from a randomized controlled trial. *Am J Med.* 1993;95(4):389-397.
72. Rehberg RS. *Classroom versus Computer-based CPR Training: A comparison of the Effectiveness of Two Instructional Methods* [dissertation]. Cypress, CA: Touro University International; 2003.
73. Childs JC, Sepples S. Clinical teaching by simulation lessons learned from a complex patient care scenario. *Nursing Education Perspectives*. 2006;27(3):154-159
74. Armstrong Medical. Nursing Annie VitalSim™. http://www.armstrongmedical.com/ami/item.cfm?sction=2&sbsection=10&category=38&itemid=2470. Accessed July 10, 2007.
75. Euliano TY. Small group teaching: clinical correlation with a human patient simulator. *Advan Physio Edu*. 25: 36-43, 2001; 1043-4046.