

association for surgical education

# The Objective Structured Clinical Examination



2<sup>nd</sup> Edition

**PREPARED BY**

**THE COMMITTEE ON TESTING AND EVALUATION**

**ASSOCIATION FOR SURGICAL EDUCATION**

**EDITORS**

**James A. McCoy, M.D.**

**Hollis W. Merrick, M.D.**

## Table of Contents

<b>EDITORS.....</b>	<b>2</b>
<b>FOREWARD TO THE FIRST EDITION .....</b>	<b>7</b>
<b>FOREWARD TO THE SECOND EDITION.....</b>	<b>8</b>
<b>THE OBJECTIVE STRUCTURED CLINICAL EXAMINATION:       PRESENT AND FUTURE .....</b>	<b>9</b>
<b>THE COMPONENTS OF THE OSCE .....</b>	<b>11</b>
REFERENCES .....	17
<b>THE ROLE OF THE OSCE IN THE ASSESSMENT OF STUDENT PERFORMANCE .....</b>	<b>18</b>
What is the role of the OSCE? .....	18
Why administer OSCEs in the 3 <sup>rd</sup> year clerkship? .....	20
Disadvantages of administering OSCEs in the junior clerkships .....	23
How we do it.....	24
Conclusions.....	25
REFERENCES.....	26
<b>IMPLEMENTATION OF THE OSCE.....</b>	<b>28</b>
Introduction .....	28
Developing the OSCE.....	28
Coordinating Committee .....	28
Other Personnel .....	29
Authoring Team.....	29
Content .....	29
Station components.....	30
Instructions to the examinee .....	30
Skill assessment checklist .....	30
Items for post encounter testing.....	30
Instructions for the standardized patient.....	31
Artifacts and equipment .....	31

Patient recruitment and training .....	31
Costs/Budgeting .....	32
Administering the OSCE .....	32
Location .....	32
Personnel.....	33
Setting up the OSCE .....	33
Examination Day .....	33
References.....	35
<b>IMPLEMENTATION AND USES OF OBJECTIVE STRUCTURED CLINICAL EXAMINATIONS (OSCE).....</b>	<b>36</b>
Organization of an OSCE .....	36
Conducting an OSCE .....	39
OSCE Results .....	39
Disadvantages/Limitations of OSCE .....	40
Applications of OSCE .....	41
References.....	42
<b>HOW TO PREPARE AND IMPLEMENT AN OBJECTIVE STRUCTURED CLINICAL EXAMINATION .....</b>	<b>43</b>
Introduction .....	43
An Example: The 1987 CSAU-OSCE.....	44
How to Develop an OSCE .....	45
The Coordinating Committee .....	45
The First Meeting.....	45
The Authoring Team.....	46
Guide for Creating a Station.....	47
The Instructions to the Candidate .....	47
The History/Physical Examination Checklist (The First Five Minutes) .....	47
The Problem-Solving Questions ( The Second Five Minutes).....	48
Instruction for the Standardized Patient.....	48
Equipment Listing for Stations .....	49

Patients for the OSCE: Real or Simulated.....	49
Follow-up Meeting .....	49
The Training of the Examiners .....	49
Choosing a Site for the OSCE .....	50
Running the OSCE .....	50
Setting up the OSCE .....	51
REFERENCES.....	53
Table 1 .....	54
<b>ADDITIONAL USES FOR OSCES .....</b>	<b>80</b>
References:.....	85
<b>EARLY FEEDBACK FOR AN OSCE .....</b>	<b>90</b>
<b>MEASUREMENT CONSIDERATIONS IN PLANNING AND EVALUATING AN OSCE.....</b>	<b>95</b>
REFERENCES.....	104

---

## Acknowledgements



*We thank the Association for Surgical Education and its founders for making this work possible. We thank the authors and member institutions for their expertise in the preparation of this manual. Our special thanks to the coordinators and administrative support staff who have processed, edited and reprocessed this manuscript. Lastly, we are grateful to the many patients who entrust us with their lives daily in practice, teaching and research. May we always strive to serve you with our very best skill, dedication, integrity and compassion.*

---

**FOREWARD TO THE FIRST EDITION**

This Manual is designed to describe some of the basic principles and concepts of the Objective Structured Clinical Examination (OSCE). We hope that it will be useful not only to the Departments of Surgery of our member institutions, but to other clinical disciplines as well.

Hollis W. Merrick, M.D.

Editor

## FOREWARD TO THE SECOND EDITION

When the first edition of this Manual was published in 1989, the Objective Structured Clinical Examination (OSCE) was a relatively new tool in the effort to obtain an accurate measurement of medical student knowledge and the effectiveness of the surgical clerkship curriculum. Eleven years later, as more and more institutions have accepted its validity and have initiated their own OSCE programs, the need remains for a basic guidebook on the principles and concepts of the OSCE. It is our combined hope that this Manual will prove helpful to all users, novice and experienced alike.

James A. McCoy, M.D.

Hollis W. Merrick, M.D.

Editors

**THE OBJECTIVE STRUCTURED CLINICAL EXAMINATION:  
PRESENT AND FUTURE***Hollis W. Merrick, M.D.**James A. McCoy, M.D.*

The advent of the Objective Structured Clinical Examination (OSCE) has offered a strikingly new and exciting way of making valid assessments of the clinical performance of medical students, residents, and fellows. Since its introduction by Dr. Harden and colleagues in 1975, the technique has gained steady and widespread acceptance around the world. Reports in the literature concerning its use come from England, Scotland, Australia, South Africa, Nigeria, the Netherlands, Canada and the United States.

The advantages of the examination are greatly apparent when one reviews the wide spectrum of clinical tests that can be incorporated into the OSCE. Such tests include multiple choice written questions, reviews of radiographs, use of models, and examination of simulated or real patients. The breadth of data that can be encompassed in this type of examination is limited only by the imagination of the examiners.

Nevertheless, there are drawbacks to the OSCE. It is cumbersome and expensive to set up, and it requires a great deal of personnel for its implementation. The expense involved in obtaining the examination site, the use of models or simulated patients, and the time of the examiners, can often be intimidating factors for those considering using an OSCE format for evaluating students, residents, and fellows.

The Objective Structured Clinical Examination (OSCE) has in some institutions become a finely tuned instrument used to evaluate clinical skills, attitudes and behaviors that are considered standards used by practitioners in the care of their patients. It is hoped that the competencies demonstrated by students, residents and fellows in the OSCE will be manifested daily in the lives of their patients as the medical profession struggles to maintain skillful, high quality, competent health care in the United States as managed care continues to erode the doctor-patient relationship.

This 2<sup>nd</sup> edition will offer different perspectives and implementation protocols from different institutions. Whereas the core components are the same, the philosophy and implementation by different institutions with varying resources are different. An innovative addition, the Early Feedback OSCE, is presented by the Medical College of Ohio. Merrick, et al. used the computer to compute the student and the residents assessments throughout the examination. Thus shortly following completion of the examination, the examinees receive their objective computer generated assessment.

The objective of this Manual is to summarize the state of the art as it exists in the United States and Canada at this point, and to make this information available to our member institutions. Many of the authors in the Manual have had extensive experience with the OSCE.

The NBME/USMLE continues on a path to introduce the OSCE as a part of the USMLE sometime in the future. This represents a massive undertaking. Even though the Royal College of Physicians and Surgeons of Canada uses the OSCE as standard to evaluate its foreign medical graduates, the number of examinees is small in comparison to the 15,000 plus American Medical Graduates. The ASE should, perhaps, help medical schools establish OSCE standards that could be approved by the NBME/USMLE and implemented by local or regional medical schools.

The OSCE does allow examiners to document competencies in clinical skills, behaviors and attitudes in an examination setting. There is no question that these objective measures should be made. What impact has this had on improving care of our patients? How do we measure commitment integrity and dedication?

Who will take the time to obtain the history, do the physical exam, generate and execute the appropriate diagnostic and treatment plans that are evidenced-based? How do we measure commitment, integrity, and dedication?

## THE COMPONENTS OF THE OSCE

*James A. McCoy, M.D.*

The OSCE (Objective Structured Clinical Examination) has reached a stage of development that allows clear recognition of key components that bring structure and organization to its construction, implementation and assessment of its performance.

The major components are:

1. The (examination) coordinating committee
2. The examination coordinator
3. Lists of skills, behaviors and attitudes to be assessed
4. Criteria for scoring the assessment (marking scheme of checklist)
5. The examinees
6. The examiners
7. Examination site
8. Examination stations
  - 8.1 Time and time allocation between stations
  - 8.2 Anatomic models for repetitive examinations (Breast, Pelvic/Rectum)
  - 8.3 Couplet Station
  - 8.4 Examination Questions
  - 8.5 Environment of Exam Station
  - 8.6 Examination Station Circuit (Figure 1.)
9. Patients Standardized or Simulated
  - 9.1 Instruction to Patients
10. Timekeeper, time clock and time signal
11. Contingency Plans
12. Assessment of Performance of the OSCE

Since its introduction into clinical medicine by Harden, et al. 1975, the OSCE has been widely used to evaluate students in virtually all areas of medicine with extension into the formative evaluation of residents. Its final form may be simply supported on a shoestring budget as described by Poenaru, et al. or as elaborate and sophisticated as administered by the Faculty of Medicine, University of Toronto (1, 2, 3, 4). It is clear, unless the major components are used as a foundation for the OSCE's structure, it will be severely flawed and its implementation fraught with difficulty and hardship. In order to develop and conduct an OSCE, appropriate clerical support through all phases of development and implementation is essential.

### **1. The Examination Coordinating Committee**

An examination coordinating committee is made up of members who are committed to the evaluative and educational process and whether appointed or volunteered must give this effort high priority in order for the OSCE to be developed and implemented. The number of members who make up this committee is not as important as the intensity of the investment of each member. It is the responsibility of the examination committee to determine the content of the examination, development and implementation. It is important that this committee has the capacity and personnel to address decisions related to reliability and validity (5).

### **2. The Examination Coordinator**

The functions of the examination coordinator (M.D. or Ph.D. educator) are the catalyst that facilitates the smooth working of the committee in developing, implementing and assessing the performance of the OSCE. For examinations conducted at different sites, a local site coordinator must work closely with the examination coordinator and the coordinating committee.

### **3. Lists of Skills, Behaviors and Attitudes to be Assessed**

The examination will measure objectively the competencies in specific areas of behavior, techniques, attitudes and decision-making strategies based on the objectives of the course or the requirement of the licensing body. The OSCE should be able to reliably assess clinical competence in history taking, physical examination, laboratory, radiographic and other data interpretation, technical and procedural skills as well as counseling and attitudinal behaviors.

### **4. Criteria for Scoring the Assessment (Marking Scheme or Checklist)**

A marking scheme or checklist is prepared for each station. Preparation of the checklist requires predetermined objective criteria that are agreed upon by the examination committee, based on faculty input. Marking scheme/checklist should be concise, unambiguous and written to contribute to the reliability of the station. The more focused the checklist, the greater the power of the station to differentiate effectively among the abilities of students. It is helpful to include instructions to the examiners at the top of each checklist/marketing scheme. The series of marking

schemes/checklist may be presented to each examinee in bound form also noting the order in which the examinee will proceed through the examination stations.

As Papp stated, “The competence of students entering the “real world” of medical practice may truly determine life or death for their patients. The development of more accurate methods of assessing clinical competence works to the benefit of the patient, the new doctor, and the medical profession in general (6)”.

## **5. The Examinees**

The examinee is the student, resident, or fellow in training or at the end of training of a prescribed course designed to teach certain clinical competencies that the examinee can use in a clinical situation to make an assessment and develop a diagnostic formulation that culminates in a therapeutic plan.

## **6. The Examiners**

Most stations will require an examiner, although some stations do not. The examiner at the station where clinical skills (history-taking, physical examination, interviewing and communication) are assessed, may be either a physician or a standardized patient. Research has shown that an acceptable level of reliability can be achieved with either a physician or standardized patient as the examiner (7,8,9). Harden recommends using examiners from a range of specialties and disciplines, for example, a dietitian at a station for nutrition. (10).

## **7. The Examination Site**

The examination site is part of a special teaching facility in some institutions. When such facilities are not available, the examination may be conducted in an outpatient facility or other space where offices are available in close proximity to each other.

## **8. Examinations Station**

The total number of stations will vary based on a function of the number of skills, behaviors and attitudinal items to be tested. For most clerkships or courses, the total will vary from 10-25. Fewer than 10 are probably inadequate and greater than 25 become from most unwieldy (11). Larger numbers of stations are located in Provinces in Canada.

### **8.1 Time Allocation and Time between Stations**

The competency being assessed in particular station will define how much time should be allotted per station. The length of time will range from 5-20 minutes. However, the time per station is constrained by the total duration of the examination. The time allocated per station should be as uniform as possible thus facilitating the smooth movement of examinees from station to station. Transit time must be built into the total time allocated for each station, e.g., a 10 minute station, 9 minutes is

allocated for the task and one minute transit time to the next station. The examiner can complete the checklist prior to the entry of the next examinee.

### **Examination station content:**

#### **8.2 Anatomic Models for Repetitive Examinations (Breast, Pelvic/Rectum)**

The skill, behavior or attitude to be tested in a station determines whether the station requires a real patient, simulated patient, laboratory data roentgenographic sticker, other clinical data, patient records, or examining equipment (stethoscope, ophthalmoscope, etc.) Chronic patients (stable) may serve well in this situation with proper training. Simulated patients who are well-trained offer reliability and consistency in the quality of their presentations. A complete material and equipment list must be generated. Specially constructed plastic models or simulations may be used, i.e. rectal or breast models, where the repetitive and intimate nature of the task makes the use of patients inappropriate. When examination equipment is a part of the station, it must be in good repair and replacement parts must be readily available.

#### **8.3 Couplet Station**

Some competencies may best be assessed by coupled or linked stations. For example, a couplet station may consist of a history-physical examination combined with a problem-solving station.

#### **8.4 Examination Questions**

Examination questions are designed to assess the ability to interpret information and critical thinking. The questions deal with diagnostic investigations, differential diagnostic and management plans.

#### **8.5 Environment of Exam Station**

The Examination Station environment should be conducive to the competency to be tested, including adjustable lighting for fundoscopic examinations and appropriate examination tables for focused physical skills assessment. Stations where auscultatory skills are being assessed should be either well insulated or in appropriately quiet areas of the examination site. Clearly marked directions leading from one station to the next should be displayed.

#### **8.6 Examination Stations Circuit Stations**

The Examination stations should be clearly marked in a logical sequence that allows easy, unimpeded transit from one station to the next.

### **9. Patient (Real) or Simulated**

A standardized patient is an individual with a health problem that is in a chronic but stable condition. Standardized or simulated patient may be used when properly trained for history and physical assessments. Simulated patients may come from the ranks of

volunteers, or acting guilds. They too, must be trained. Whether real or simulated, more training is required for patients used in history taking than for patients used for physical examinations.

- 9.1 Detailed instruction package is provided for both the standardized and simulated patient. The instructions describe how the patient responds to historical questions and physical exam, as well as how the patient should dress.

## **10. Timekeeper, Time Clock and Time Signal**

Appropriate personnel for the position of official timekeeper and exam facilitators need to be identified and properly instructed. A well functioning time clock and time signal are critical. The time signal should be unambiguous and clearly audible throughout the examination site. (11). One support person per three stations is recommended.

## **11. Contingency Plans**

A contingency plan includes reserve standardized patients who are trained to assume a number of roles, and a patient trainer who circulates to deal with any patient problems that arise. A number of reserve stations should be available. A contingency plan must be developed for students who must leave the exam when the situation arises.

## **12. Assessment of the Performance of the OSCE**

It is the responsibility of the Examination Committee to address the numerous measurement issues related to the OSCE. DaRosa cites five key areas that should be addressed (12):

1. The OSCE should be tested for appropriate measurement characteristics such as validity, reliability, feasibility and credibility.
2. The different types of validity for which an OSCE can be tested include predictive validity, concurrent validity and content validity (8).
3. A valid OSCE station measures what it was designed to measure. A reliable station measures it consistently (13).
4. Item analysis should be completed for an OSCE to provide indications concerning the difficulty of each station in relation to the overall exam and how each station discriminates among various levels of performance.
5. Grading can be based on a criterion-referenced system, norm-referenced system, or a combination of both. The Examination committee needs to decide in advance which system best meets its fundamental purposes for the exam.

**REFERENCES**

1. Harden, R.M. Stevenson, M. Downie, W.W, Wilson, G.M., Assessment of Clinical Competence Using Objective structured Examinations. British Medical Journal 1 447-451, 1975.
2. Poenaru, D., Morales, D., Richards, A., O'Connor, H.M. Running on Objective Structured Clinical Examination on a Shoestring Budget. American Journal of Surgery, 173 538-541, 1997.
3. Sloan, DA, Strodel, WE, Blue AV, Felts JL Schwartz RW, Dannelly, MB. The Use of Objective Structured Clinical Examination (OSCE) for Evaluation and Instruction in Graduate Medical Education. Journal of Surgery Resident 63 (1) 225-30. '96
4. Cohen, R. MacInnes, A, Finkelstein J. Rothman AI. Robb BA How to Prepare and Implement An Objective Structured Clinical Examination. OSCE Ed. Merrick HW, Prater MA pp 24-40 1989.
5. Sachdeva, AK. Implementation and Uses of Objective Structured Clinical Examination from Merrick HW, and Prater, MA. Ed. OSCE pp, 11-23, 1989.
6. Papp, KK, The Role of the OSCE in the Assessment of Clinical Competence. OSCE, ed. Merrick HW. Prater MA 4-10 1989.
7. Stillman, P.L., Swanson, D.B., Smec, S., et al. Assessing Clinical Skills of Residents with Standardized Patients. Annals of Internal Medicine 105: 762-771, 1986.
8. Robb, K.V., Rothman, A.I. The Assessment of Clinical Skills in General Medical Residents: Comparison of the Objective Structured Clinical Examination with a Conventional Oral Examination. Annals of the Royal College of Physicians and Surgeons in Canada 18:235-238. 1985.
9. Amiel, G.E., Tann, M. Krausz, M.M., Bitterman, A. Cohen, R., Incurring Examiner Involvement in an Objective Structured Clinical Examination by Intergrating A Structured Oral Examination, American Journal of Surgery, 173 546-549, 1997.
10. Harden, RM. Twelve Tips for Organizing an Objective Structured Clinical Examinations. (OSCE) Medical Teacher 12 #314, 259-264, 1990.
11. Selby, C., Osman, L., Davis, M., Lee, M. Set Up and Run and Objective Structured Clinical Exam. British Medical Journal 310 1187-1190, 1995.
12. DaRosa, FA Measurement Considerations and Planning and Evaluating An OSCE. OSCE, Ed. Merrick, HW. Prater MA p 71-83. 1989.
13. Skinner, B.D., Newton W.P., Curtis P. The Value of an OSCE in a Family Practice Residency. Academic Medicine 72 #8 722-724, 1997.

## THE ROLE OF THE OSCE IN THE ASSESSMENT OF STUDENT PERFORMANCE

Sally L. Hodder, M.D.

Klara K. Papp, Ph.D.

### What is the role of the OSCE?

The importance of having multiple measures to evaluate the complex multi-dimensional facets of clinical performance has long been recognized. The multiple-choice test alone is inadequate. The following excerpt from the “Flexner Report” makes this point (Flexner, 1910):

*“There is only one sort of licensing test that is significant, viz., a test that ascertains the practical ability of the student confronting a concrete case to collect all relevant data and to suggest the positive procedure applicable to the conditions disclosed. A written examination may have some incident value; it does not touch the heart of the matter”.*

The corollary of this point in surgical terms follows: “...those who evaluate students must have adequate biopsies from each competency domain. Paper-and-pencil examinations such as the surgery part of the National Board of Medical Examiners (USMLE Step 2) subtest make no pretense of evaluating all aspects of students competence in surgery. Evaluation of the overall competence of students as part of the process of assigning a clerkship grade is deficient if only knowledge is measured, but is equally deficient if these other important attributes [motor skills including coordination, interpersonal skills, drive, and honesty] are not evaluated (2).”

The Objective Clinical Examination (OSCE) is designed to assess clinical competence. It was developed in Dundee, Scotland in the early 1970’s (3). The OSCE is a group of different stations (typically 20 to 30) through which each student must rotate. The competencies to be tested are distributed among the different stations. Stations are characteristically of three types: a) history-taking stations, b) physical examination stations, and c) short answer stations. At each station, the student is given a very brief (one to two sentence) clinical vignette and instructed to perform a specific task. For example, the student may be asked to take the history of a young woman presenting with a breast lump. In a subsequent station, the student may be requested to perform the breast examination on a standardized patient (or on the provided breast model). A list of questions to be asked (in the case of the physical examination station) constitute the “correct answers” and are specified by the faculty prior to the examination. These examinations are carefully timed, with 4 to 5 minutes usually allotted for the performance of each station. A one minute feedback sessions during which the faculty member gives the student formative instruction on their

performance is optional. Students are generally given less than a minute to change stations.

The purpose of this chapter is to review the arguments for and against including the OSCE in the evaluation of third-year surgical clerkship students. At the same time, we will try to determine whether this form of assessment, growing in popularity over the past three decades, is merely a passing trend (4) or a method of medical student evaluation that is here to stay. We would like the reader to know that our discourse is grounded in experience; our department administers the OSCE examination regularly after each clerkship and had done so for the past twelve years. We describe some of the practical issues involved.

## Why administer OSCEs in the 3<sup>rd</sup> year clerkship?

The OSCE offers many advantages to medical students as well as to faculty who direct educational programs for students. Five reasons for administering OSCEs to junior clerkship students are discussed.

⇒ *Provides a opportunity to test a student's ability to integrate knowledge, clinical skills, and communication with the patient*

Practicing physicians integrate medical knowledge with clinical skills and effective patient communication to solve clinical problems (5). Rarely, however, is this integrative ability ever objectively assessed in the third-year medical student. Though student-patient interactions may be observed by faculty members serving as attending physicians in the inpatient setting or as preceptors in the ambulatory venue, those observations frequently lack an agreed upon objective list of items on which the student is checked. OSCEs provide a uniform objective upon which students are evaluated. For example, a student is asked to elicit a history from a patient with chronic cough. The student's ability to elicit an appropriate history is determined by the number of questions, from a pre-determined faculty generated list of items; the student actually asks the patient. The more questions that the student omits, the poorer his/her performance. Student performance may also be assessed by a list of desirable physician behaviors to be exhibited. The student's capacity to communicate clearly, make the appropriate introductions, put the patient at ease, establish eye contact, and ask sensitive relevant questions in a compassionate manner may also be assessed in an OSCE station in which a patient history is being elicited.

⇒ *Provides the faculty with an assessment tool that is custom-fit to the goals of a specific education program*

In the not-too-distant past, a student's knowledge of a discipline was largely dependent on the clinical material with which the student came into contact. A core-curriculum now exists for many of the disciplines through which junior students are required to rotate. However, within a core-curriculum, there exists considerable liberty for topic choice and emphasis. Even within a single curriculum item, there may be many potential competencies for students to master. For example, breast cancer is a ubiquitous disease that third-year students could be expected to develop competence. Eliciting an appropriate history of a breast lump, performing an accurate breast examination, and informing a patient that she has a diagnosis of breast cancer with various treatment options are all examples of competencies that third-year students may be asked to master in regards to breast cancer. Mastery of any one of these competencies requires demonstration of successful completion of multiple elements. Administration of a OSCE requires the faculty to define those domains in which it is crucial for a third-year student to be competent. Moreover, within each domain, a list of items to be performed must be generated. This process has been termed examination blueprint development (6). As Page, et al describe, an examination blueprint may seek to distribute various tasks among various body systems (cardiac, endocrine), various venues (ambulatory vs. inpatient), or various

skill categories (eliciting a patient history or performance of physical examination). By development of an examination blueprint, faculty have the latitude to determine the distribution of tasks which best fits with the goals of their specific educational program. After determination of the distribution of tasks, the faculty must then develop for each task, and objective list of items by which the student will be evaluated. Essential items may be weighted to be worth more points than items considered preferable but not essential (7). OSCE examination development presents an excellent way for faculty to focus the goals of an educational program by clearly defining the domains of clinical competence important for students to master.

Though many important skills such as putting the patient at ease or asking sensitive questions in a compassionate manner are not directly taught, they are essential to the practicing clinician. An OSCE provides the opportunity to direct the students' attention to mastering skills that may otherwise receive little, if any, attention elsewhere in the educational program. A major motivation to develop OSCEs provided not by the actual scores obtained using them, but because of its influence on what is taught and what is learned. The assessment tail wags the curriculum dog, is a well known saying in medical education. Testing procedures guide students year-to-year to focus on what faculty consider important and essential to learn throughout a course of study.

⇒ *Renders an occasion for individualized instruction and feedback*

The third-year of medical school is a time of intense learning of skills that will be used throughout a physician's lifetime. During this formative period, it is crucial that proper methods be learned to avoid the establishment of incorrect techniques which pave the way for a professional lifetime of bad habits that may compromise clinical care. The OSCE provides an opportunity to reinforce properly performed skills and to correct improper skills at a time when the student is still in a formative phase. For example, OSCE stations that require students to perform a particular part of the physical examination such as the breast examination, provide the opportunity to assure that the student is using an appropriate technique and performing complete examination (7). A student's error such as failure to examine the entire breast in an orderly fashion may be remedied immediately. A strategy that we have successfully used to integrate individualized instruction into the examination is to provide one minute of feedback at each station to immediately follow the five-minute period that students are given to complete an assigned task. Student welcome the feedback and report this experience to be an outstanding learning experience (Unpublished data). Integrating feedback into the OSCE also provides the faculty evaluators at each station with the opportunity to be actively involved. Amiel, et al recently reported a "high level of dissatisfaction" among faculty examiners due to their passive role during examinations utilizing standardized patients (8). The same authors report improved examiner satisfaction when they were given a more active role during the examination. By establishing a feedback time at each OSCE station, faculty evaluators are active participants who administer individualized instruction while performing the summative evaluative aspect of this examination. Concerns arise concerning the security of an examination when feedback on student performance is given and identical stations are reused. However, conflicting results have been published on the effect upon student OSCE scores of reusing standardized patient

based examinations (6, 9, 10). While these concerns are legitimate, their significance is unclear. It is our conclusion that the positive effects of giving individualized student feedback during OSCEs currently outweigh the uncertain effect on security.

⇒ *Offers an additional parameter by which to evaluate student performance*

A student's core-clerkship evaluations are an important determinant of where a student will be accepted for post-graduate training and occasionally will even effect a student's career choice. It is important to have multiple data points from which to determine a student's final grade for several reasons. Given the "high-stakes" of final clerkship grades, the presence of multiple data points decreases the likelihood that a student's grade will be adversely affected by a spurious evaluation that may not accurately reflect the student's true performance. Furthermore, student anxiety may be diminished by eliminating schemes of evaluation where performance on only one or two occasions may determine the final grade. Much clerkship use clinical evaluations completed by attending physicians and residents who have worked with the student in the clinical setting. Written multiple choice questions are used by many programs to determine a student's final grade. The OSCE provides an additional useful parameter by which to evaluate students.

⇒ *Provides unique programmatic evaluation*

OSCEs provide unique opportunities for programmatic evaluation in several ways. Student group performance on a given station or group of stations may indicate weaknesses or strengths of the educational program. For example, poor overall student performance on a specific station such as the abdominal examination may indicate to the Clerkship Director the need to add specific instructional sessions to improve this skill.

The effectiveness of specific educational innovations added to a program may also be assessed using the OSCE. For example, one-half the class may receive additional didactic sessions while a control group proceeds through the conventional clerkship. An OSCE given near the completion of the clerkship may contain one or more stations seeking to assess students' knowledge, skills, and/or attitudes relevant to the educational innovation. Data may be generated that not only assesses the global effect of the innovation, but may assess the innovation's effect on certain aspects of the content area. This may be accomplished by using item list analysis to dissect student performance relative to specific points of knowledge, precise skills, or certain attitudes.

Not only may the effectiveness of innovative educational programs be assessed, but one may evaluate the effect of various methods of educational delivery. The OSCE has been used to demonstrate superior performance on various asthma related tasks by students having completed a multi-disciplinary asthma clinic as compared to students given didactic sessions but nonspecific clinic experience (11). Clearly the OSCE is a valuable resource for effective programmatic evaluation.

## **Disadvantages of administering OSCEs in the junior clerkships**

⇒ *Provides assessment of case-specific skills, knowledge, and/or attitudes*

Though examinees are tested in a clinical situation using “real” or “scripted” patients, the domains that are assessed are often specific to a certain clinical problem rather than to a skill that may be generalized across clinical problems. For example, if one purpose of the test is to measure students’ physical examination skills, it should be possible to develop a list of skills students need to master and situations in which those skills are required with sampling from among those situations. Summing across those stations should allow faculty to make inferences about students’ physical examination skills in those contexts. Unfortunately, it is not that simple. Physical examination skills are multifaceted and, depending upon the sampling plan, markedly different scores could be obtained for individual examinees.

The research in this area is consistent. No evidence for underlying fundamental competence, such as history taking, physical examinations stations, would be well advised to concentrate efforts on improvement of physical examination skills in the cases examined. We have no way of knowing whether the student’s examination skills in situations not presented on the OSCE are adequate or not, however, we do know that the student’s performance on those stations tested was below his/her performance level on other parts of the OSCE. This information is an important piece of formative evaluation for students.

⇒ *Development and administration are time consuming and costly.*

Developing OSCEs is time-intensive and costly. In a world of shrinking resources, faculty must consider the cost and energy invested into this examination whose cost has been estimated to range from \$21 to \$1000 per examinee (12). Even on a “shoestring budget”, OSCEs were estimated to cost \$50 - \$70 per student (14). While this may appear to be “expensive”, OSCE costs are certainly within the same ballpark as standardized multiple choice tests that are administered in many departments. Moreover, there are those intangible benefits such as active faculty involvement that may provide many added values (e.g., focusing the curriculum, providing individualized student instruction) to the educational program.

⇒ *Offers opportunity for compromised test security*

Legitimate concerns exist about the security of an examination in which stations are used repeatedly. Providing students feedback on the answers or skill demonstration that is sought intensifies those concerns.

Several studies have failed to show higher scores on reuse of performance-based assessments. Stillman et al examined the effect on the mean scores of successive testing of multiple student/resident populations (10). Though these data suggest that security is not significantly compromised by administering performance-based examinations at different times, the authors discuss that the low-risk stakes of their examinations are well as the dispersal of test takers may have resulted in less dissemination of the test questions. Colliver et al were also unable to demonstrate

significant increases in scores of an examination administered at different times (13). Using the Wilcoxon matched-pairs sign-test, Cohen et al found an increase in mean scores of reused OSCE stations (9). The authors suggest that candidates' scores may improve as a result of concentration of their efforts on subjects repeatedly used in the examination rather than of sharing the answer key. These authors further conclude, "Inclusion of a limited number of stations used repeatedly over a number of years appears to have minimal impact on total scores on the examination (9)".

The degree if at all, to which reusing OSCE stations compromises the integrity of the examination is not established. Repeated reuse of an OSCE station should prompt longitudinal analysis of the scores obtained on the station used on repeated occasions should raise the possibility of compromised security and thought should be given to retiring the station. However, improved scores may also indicate, as proposed by Cohen et al (9), that students are focusing their efforts on topics present in past examinations. This may be a desirable effect for a limited number of stations.

In our experience, reusing short answer stations (rather than history taking or physical examination stations) has raised the greatest concern for security of the examination. Therefore, answers to short answer stations are neither discussed with the student at the time of the practice.

Student performance information generated by OSCE examinations should not be used in a vacuum. As we have discussed, multiple evaluation parameters are preferable and information from clinical evaluations of the student while in the clinic and operation room need to be used with the OSCE data to generate a legitimate profile of student performance.

## **How we do it**

Scores at each station are derived from faculty-scored 3-point rating scales (1=full credit, .25=partial credit, 0=no credit). The maximum number of points achievable at each station may vary considerably. On our last OSCE (22 stations), the maximum total points ranged from a high of 46 to a low of 5.

We calculate a % correct raw score for each student on each station. The scores are then standardized to mean=500, standard deviation=100. Overall any students who scores below two standard deviations of the mean of the group is required to retake the OSCE exam at a subsequent administration. An individualized student performance sheet as shown in Table I is sent to each student following the OSCE.

The standardized OSCE score is weighted with other clerkship measures in the calculation of the students final grade as follows:

	<u>% weight</u>
NBME subject examination	.25
OSCE examination	.25
Inpatient evaluation	.40
Ambulatory evaluation	.10

## Conclusions

Use of performance-based testing appears to be expanding. The NBME's Strategic Plan calls for the development of a performance-based test of clinical skills using Standardized Patients. The project is beyond the test development phase and into the implementation planning (NBME Annual Report, 1996, p. 43). In Canada, successful completion of a performance-based examination is already a requirement for licensure (15). The Liaison Committee on standards of medical schools in 1991 (16). If performance-based assessment is merely a new fashion and latest trend in medical education, it has many prominent advocates. Our guess is that its success may be partially explained by Abraham Flexner's comment quoted in the beginning of this chapter; the OSCE ascertains the practical ability of the student confronting a concrete case. It touches the heart of the matter.

We are not certain whether the institution of OSCEs during medical school will result in graduating better physicians. However, institution of OSCEs into educational programs provides unique opportunities for student education and evaluation as well as for programmatic assessment. Clerkship directors are well advised to seriously consider the advantages, disadvantages, and feasibility of introducing performance-based assessment into their educational programs.

## REFERENCES

1. Flexner A. 1910 quoted in Swanson DB, Norman GR, Linn RL. Performance assessment: Lessons from the health professions. Educational Researcher 24: 5-11, 1995.
2. Chase RA. Paper and pencil examinations – What they can and cannot do. Surgery 94: 771-772, 1981.
3. Harden RM, Stevenson M, Downie WW. Assessment of clinical competence using objective structured clinical examinations. British Medical Journal 2975 1: 447-449.
4. Mavis BE, Henry RC, Ogle KS, Hoppe RB. The Emperor's New Clothes: The OSCE Reassessed. Academic Medicine71: 447-453, 1996.
5. Sloan DA, Donnelly MB, Johnson SB, Schwartz RW, Strodel WE. Assessing surgical residents' and medical students' interpersonal skills. Journal of Surgical Research 57: 613-618,1994a.
6. Page G, Bordage G, and Allen T. Developing Key-Feature Problems and Examinations to Assess Clinical-Decision-making Skills. Academic Medicine 70: 194-201, 1995.
7. Sloan DA, Donnelly MB, Johnson SB, Schwartz RW, Munch LC, Wells MD, Johnson SB, and Strodel WE. Assessing Medical Students' and Surgery Residents' Clinical Competence in Problem Solving in Surgical Oncology. Annals of Surgical Oncology 1:204-212, 1994b.
8. Amiel GE, Tann M, Krausz MM, Bitterman A, and Cohen R. Increasing Examiner Involvement in an Objective Structured Clinical Examination by Integrating a Structured Oral Examination. American Journal of Surgery 173: 546-549, 1997.
9. Cohen R, Rothman A, Ross J, and Poldre P. Security Issues in Standardized-Patient Examinations. Academic Medicine 68(supp): S73-S75, 1993.
10. Stillman PL, Haley HA, Sutnick AI, Philbin MM, Smith SR, O Donnell J, and Pohl H. Is test security an issue in a multi-station clinical assessment? - A preliminary study. Academic Medicine 66 (supp.): S25-S27, 1991.
11. Gilbert IA, Aucott J, Hull AL, Hodder SL. Traditional Internal Medicine Clerkship Training Does Not Prepare Students in Practical Issues of Asthma Diagnosis and Management. American Journal Respiratory Critical Care Medicine153: A277, 1996.
12. Carpenter JL. Cost analysis of Objective Structured Clinical Examinations. Academic Medicine 70: 828-833, 1995.

13. Colliver JA, Barrows HS, Vu NV, Verhulst SJ, Mast TA, Travis TA. Test Security in Examinations that use Standardized Patient Examinations. Academic Medicine 66: 279-282, 1991.
14. Poenaru D, Morales D, Richards A, O'Connor M. Running an Objective Structured Clinical Examination on a shoestring budget. American Journal of Surgery 173: 538-541, 1997.
15. Reznick RK, Blackmore D, Cohen R, et al. An objective structured clinical examination for the licentiate of the Medical Council of Canada: from research to reality. Academic Medicine 68 (suppl): S4-S6, 1993.
16. Liaison Committee of Medical Education. Functions and Structure of a Medical School. Chicago, IL: Council on Medical Education, American Medical Association, 1991.

## IMPLEMENTATION OF THE OSCE

Walter E. Pofahl, M.D.

David A. Sloan, M.D.

### Introduction

*The objective structured clinical examination (OSCE) is a valuable tool for assessing skills that are not easily evaluated by other testing methods. Uses of this examination range from small intradepartmental evaluations to large-scale examinations required for licensure. The purpose of this chapter is to provide the reader with the “nuts and bolts” of developing and administering an OSCE. This chapter is directed toward educators planning the initiation of an OSCE at their institution. This chapter will orient the reader to the steps necessary to successfully develop an OSCE and the logistics associated with administering the examination.*

### Developing the OSCE

*Administrative Structure*

*Coordinator*

The key to a successful OSCE is careful planning. This is especially true for the initial administration of the examination. While decisions such as the content and scope of the examination require input from a committee, there must be a single leader to coordinate the development and implementation of the examination. The coordinator is responsible for overseeing the development, organization, administration, and grading of the examination. In the case of multisite examinations, there should also be local coordinators available for each individual site. The coordinator should be identified as early as possible in the planning of an OSCE. Ideally, this person should have experience with this examination, especially if the OSCE is being administered for the first time.

### Coordinating Committee

Because of the time and energy required, a committee should be formed to develop and implement the examination. At a minimum, the committee should consist of the coordinator, a small number of faculty familiar with the curriculum and objectives, the patient coordinator/trainer, and an educator familiar with performance based testing. This group determines the content of the examination and develops

appropriate stations to test this content. At the initial committee meetings. A timeline for the development of the examination should be formulated. A recommended schedule follows: 3-6 months prior: need for an OSCE identified; coordinator appointed; committee formulated. 2-3 months prior: content finalized; stations developed; standardized patients recruited 1 month prior: stations finalized; artifacts obtained; standardized patients trained 7 days prior: walk-through; standardized patients confirmed; faculty confirmed.

### **Other Personnel**

An OSCE requires considerable clerical support for successful development and implementation. It has been estimated that approximately 80 hours of secretarial time is required to organize a large scale OSCE for the first time in an institution [8]. The majority of the time is spent on the development of the examination. Other personnel include trainers, standardized patients, and additional support personnel used during OSCE administration. The roles of these personnel will be discussed in the appropriate sections of this chapter.

### **Authoring Team**

When the examination content has been decided, the task of developing stations can begin. This should be handled by a larger group of physicians who are familiar with the curriculum and its objectives. Station authors should be provided with a detailed explanation of the OSCE format, the examination objectives, and guidelines for authoring a station. As individual stations are developed, these should be reviewed, edited, and accepted by the coordinating committee.

### **Content**

The content of the examination is determined by the coordinating committee based on the curriculum, course objectives, and examination goals. If the OSCE is to be used as a reliable tool to assess clinical competence, a broad sampling of curricular material encompassing a variety of skills (history-taking, physical examination, laboratory data interpretation, problem-solving, procedural, counseling, attitudinal) is absolutely necessary [8]. To assure this broad sampling there must be an adequate number of stations of appropriate duration. In general, there should be no more than 25 stations of 5 to 20 minutes in length. In order to ensure smooth flow, all stations should be of the same duration. Stations lasting approximately 10 minutes are usually the best compromise. The time required to move from station to station must be factored into the station length and duration of the examination. For example, with 10 minutes stations, the actual encounter lasts 9 minutes while one minute is allotted to changeovers between stations. This allows the examinee time to move to the next station and read the instructions regarding the task to be performed at each station. This additional time also allows examiners to complete their checklists prior to entry of the next examinee.

### **Station components**

The following components must be developed by the station authors for each station [2]:

An instructions sheet for the examinee.

A checklist for the assessment of the skill being examined at that station.

Items (including answers) for any post encounter testing.

A detailed patient profile for the standardized patient.

A list of the equipment and artifacts required at the station.

### **Instructions to the examinee**

Complete instructions must be written for the examinee, stating clearly and concisely what skill the examinee is expected to demonstrate at that station. This is typically in the form of a clinical vignette. For example: a 54 year old lady presents to your office with a breast mass. Obtain a history. Another option for this station would be to ask the student to perform a breast examination on the same patient.

### **Skill assessment checklist**

To evaluate the history taking, physical examination or interviewing skills of the students in those stations where patients are utilized, the examiner will use a checklist. Checklists can contain specific items (making whether or not they were performed) and/or more subjective global ratings. The coordinating committee must decide the format so that there is uniformity throughout the examination.

### **Items for post encounter testing**

Combining two stations is typically referred to as a couplet. Typically the first station tests a clinical skill, while the second station (frequently referred to as a post encounter probe) tests data synthesis. For example, at the first station the student might be asked to obtain a history from a patient with a breast mass. Possibilities for the post encounter probe (i.e. the second station of the couplet) include interpreting mammograms or photomicrographs, outlining a diagnostic plan, or writing a note outlining the history with a differential diagnosis. When the station is being developed, correct answers to the post encounter test must be determined and a scoring method defined.

### **Instructions for the standardized patient**

A detailed patient profile, understandable to a layperson presenting the role, should be written (typically in the form of a clinical vignette). This should contain detailed instructions on how the patient should dress and act during his/her interactions with the examinees. All of the historical information that the patient may be required to provide must be documented. Similarly, a detailed description of examination findings to be elicited must be provided.

### **Artifacts and equipment**

Artifacts such as radiographs or pathology slides enhance the experience of the examination. These are most useful in post encounter testing. Separate stations that require interpretations of anatomic drawings, radiographs, or photomicrographs can also be used. In obtaining these artifacts, there must be strict attention to patient confidentiality. A list of other equipment, such as blood pressure cuffs, otoscopes, and ophthalmoscopes, must be provided by the station authors.

### **Patient recruitment and training**

Because a goal of the OSCE is to simulate an actual clinical encounter as realistically as possible, the use of actual and simulated patients is an essential component to the examination. Although real patients with acute problems do not lend themselves to this type of examination, those with chronic stable findings (fundoscopic changes, goiters, adventitious pulmonary sounds, cardiac murmurs, abdominal organomegaly, ostomies, skin changes, deformities) can be used very effectively. These patients are best obtained through the assistance and cooperation of their primary physicians. Standardized patients have several distinct advantages over real patients. They can be carefully controlled and their history is more reproducible [5]. Also, their simulation of physical findings can be standardized. These patients are especially useful in the evaluation of skills where participation of real patients would cause them undue distress or embarrassment. Using standardized patients, the level of difficulty of the examination system can be more readily controlled and their use during subsequent examinations can allow accurate comparison between the standards of performance of different students [4]. Standardized Patients can also be trained to evaluate firsthand and provide feedback on such issues as display of professionalism, respect, and gentleness of technique [8]. Standardized patients can be recruited from local actors or others with an interest in medical education.

The first step in training standardized patients is the development of the patient profile. This is the blueprint for the patient interaction with the examinees. As previously noted, this must document every aspect of the interaction such as patient dress, position, behavior, mannerisms, etc. With the profile as a template, the patient coordinator works with the patient to develop a standardized, reproducible clinical

encounter. When a patient has been adequately trained for a particular station, the portrayal can be videotaped to assist in the training of subsequent patients for the station. In general, training a patient to present a history requires two to five hours, while training a patient for a physical examination station requires five to ten hours of training.

## **Costs/Budgeting**

Costs associated with developing and implementing an OSCE can vary widely. There are both indirect (hidden) and direct costs associated with the examination. Variables associated with these costs include the magnitude of the examination, its frequency, the use of standardized patients, the use of volunteer faculty. Estimates for the costs of the examination vary from \$28 to over \$1000 per student [1,3,6,7]. The overwhelming source of expense for developing and administering these examinations is personnel expenses. This can take the form of indirect or direct costs. Cusimano, et al [3] found that development and administration of a six station OSCE to 40 students required a total of 327.5 hours. Approximately 50% of this time was during the development phase of the examination. It is reasonable to assume that academic faculty will not be directly reimbursed for their participation in the development and administration of an OSCE at their own institution. Similarly, secretarial support during normal hours will not require additional direct expenditures. However, the loss of clinical revenue represents a true cost and is especially important as practice income is being monitored closely in academic faculty. The other source of direct expenditure for personnel costs is in the use of standardized patients. The hourly expense for the use of standardized patients, excluding training, ranges from \$7.50 to \$50.00 depending on the level of interaction with the examinee (obtaining a history vs. physical examination) [9]. Expenditures for artifacts should be only a fraction of overall expenditures. Aside from the indirect costs associated with identifying and locating the appropriate materials, there are costs associated with reproducing the material. There should be no direct costs associated with the use of clinic space for administration of the examination. Post examination clean up will require direct expenditures only if janitorial services are required.

## **Administering the OSCE**

### **Location**

Ideally the examination should recreate a clinical encounter in as realistic of a fashion as possible. The examination should be administered in an area comfortable for patients, examinees, and supervisors. Outpatient clinics are the most effective locations for several reasons. Most are planned to allow smooth flow from station to station. Furthermore, there are provisions for patient privacy. These same provisions also contribute to examination security. Amenities such as rest rooms and handicap access are available. The main disadvantage to using outpatient clinics is the need to schedule the examination so that it does not conflict with normal patient care activities. Thus most examinations that utilize clinic facilities must occur during the evening hours or on weekends. When planning an OSCE, the testing area should be diagrammed, stations noted and flow patterns shown. The layout should be such that

the bell or buzzer used to indicate the time to switch stations can be heard throughout the examination area.

## Personnel

In addition to the faculty and staff required to develop the OSCE, other supplementary personnel are required to administer the examination. Assistance is necessary for setting up the examination (distributing materials, setting up individual stations,), during the examination (collecting answer sheets, attending to needs of standardized patients, examiners, and examinees), and following the examination (cleanup, collating checklists and answer sheets). In many cases, these tasks can be consolidated. The patient coordinator should be available to deal with any patient problems. Depending on the duration of the examination, more than one patient may be required for each station.

## Setting up the OSCE

The week prior to the examination, detailed instructions should be sent to all examiners, examinees, and patients. The appropriate number of copies of all examination materials including checklists, examinee response sheets and instruction sheets must be made. Equipment and artifacts required for the examination must be obtained at this time. All personnel participating in the examination should be sent reminders and equipment items, as well as examination materials, must be rechecked [8]. Several additional stations that do not require standardized patients should be developed and available in case of difficulties with other stations.

## Examination Day

About 1 hour prior to commencement of the examination, each station should be set up with the appropriate equipment and examination materials. Instructions defining the student task at each station are left outside the room or at an appropriate place inside. Each station is numbered and the buzzer or bell checked. The movement from station to station should be clearly marked for the examinees. Examiners, examinees and patients are asked to arrive at the examination site at least 30 minutes prior to the examination. Orientation sessions are given to all three groups separately. Prior dissemination of detailed information relating to this examination greatly facilitates this process [8]. When all examination materials are in place and patients as well as examiners ready, examinees are asked to go to their pre-assigned stations. They are told to wait in front of their respective stations until the sound of the buzzer or bell indicating they may begin the station. At each subsequent sound of the buzzer or bell, examinees move to the next station. The process stops only when students have rotated through all stations. If the first station assigned to an examinee is based on information that the examinee should have obtained at the previous station, the examinee does not start the examination until the next rotation of stations. This

person will finish the examination at the first assigned station, a little later than the other examinees. To avoid cueing the examinees, questions or response sheets for any post encounter testing should be provided at a separate station following the station where the specified task has been carried out. If a test station is followed by a feedback station the examiner moves with the examinee to the feedback station and another examiner takes his/her place in the test station. When this latter examiner moves, in turn, to the feedback station, the former goes back to the test station and the cycle continues [8]. Patients, whether real or simulated, may be able to tolerate only a certain number of physical examinations. Hence, at appropriate intervals, the assistants may be required to relieve them by bringing in other patients with similar findings. This change can be accomplished smoothly during the time examinees are moving from one station to the next. It is also the responsibility of these assistants to attend to other needs of patients while the examination is in process [8]. At the completion of the examination, checklists and answer sheets to post encounter-testing stations are collected and collated. Grades are determined based on a predetermined scoring method.

## References

Carpenter JL. Cost Analysis Objective Structured Clinical Examinations. Academic Medicine 70: 828-833, 1995.

Cohen, R., MacInnes, A., Finkelstein, J., Rothman, A.I., and Robb, A. How to Prepare and Implement an Objective Structured Clinical Examination. In: OSCE Manual: 24-44, 1989. ASE Educational Clearinghouse.

Cusimano MD, Cohen R, Tucker W, Murnaghan J., Kodama R, and Reznick R. A Comparative Analysis Of The Costs Of Administration Of An OSCE (Objective Structured Clinical Examination). Academic Medicine 69:571-576, 1994.

Harden RM. Gleeson FA. Assessment Of Clinical Competence Using An Objective Structured Clinical Examination. Medical Education, 13:41-54, 1979.

Harden RM, Stevenson M, Downie WW, and Wilson GM. Assessment Of Clinical Competence Using Objective Structured Examination. British Medical Journal 1: 451, 1975.

Hasle JL, Anderson DS, and Szerlip HM. Analysis Of The Costs And Benefits Of Using Standardized Patients To Help Teach Physical Diagnosis. Academic Medicine 69: 567-570, 1994.

Poenaru D, Morales D, Richards A, and O'Connor MH. Running An Objective Structured Clinical Examination On A Shoestring Budget. American Journal of Surgery 173: 538-541, 1997.

Sachdeva, A.K. Implementation And Uses Of Objective Structured Clinical Examinations (OSCE). In: OSCE Manual: 11-23, 1989. ASE Educational Clearinghouse.

Stillman PL. Technical issues: logistics. AAMC Consensus Conference on the Use of Standardized Patients in the Teaching and Evaluation of Clinical Skills. Academic Medicine 68(6): 464-470, 1993.

## IMPLEMENTATION AND USES OF OBJECTIVE STRUCTURED CLINICAL EXAMINATIONS (OSCE)

Ajit K. Sachdeva, M.D., F.R.C.S.C., F.A.C.S.

*The Objective Structured Clinical Examination (OSCE) format can be used to test process as well as product. Prior to implementation of an OSCE, it is important to determine what clinical skills one wants to test and whether OSCE is the best way to test those particular skills. Once it has been decided to implement an OSCE, careful organization and execution of a series of well planned steps are absolutely essential for successful conduction of the examination. Even the most minute detail must be carefully considered to reduce the risk of a catastrophe occurring on the day of the examination. Adequate time, as well as resources, must be allocated for the different stages of implementation of the examination (1).*

### Organization of an OSCE

An examination coordinator should be appointed approximately 6-8 weeks prior to the scheduled examination. This person plays a pivotal role throughout the planning and conduction of the examination. The examination coordinator should be assisted by a committee of key faculty members who are familiar with the curriculum, course objectives and goals of the examination. If the examination is to be conducted at different sites, each site must have a local coordinator who will work closely with the examination coordinator responsible for the entire examination.

It is the responsibility of the examination committee to determine the content of the examination and select appropriate stations for the examination. If the OSCE is to be used as a reliable tool to assess clinical competence, a broad sampling of curricular material encompassing a variety of skills (history-taking, physical examination, laboratory data interpretation, problem-solving, procedural, counseling, attitudinal) is absolutely essential. Sequential stations may be set up dealing with the same clinical problem; a history-taking station may be followed by a physical examination station and then a laboratory/x-ray interpretation station. One may, however, limit the scope of the examination to a specific area, depending upon the course objectives and time allocation for the examination.

Once the examination content has been determined, the individual stations must be selected. The total number of stations can vary from 10 to 25, depending upon the different items to be tested, as well as the number of candidates. Handling more than 25 stations at any one time can become a logistical nightmare. If large numbers of candidates are to be tested, the examination may need to be repeated or a second examination site set up with identical stations, allowing two groups of examinees to be tested simultaneously. Time allocated for each station can vary from 5-20 minutes, depending upon the skill to be tested and the task the examinee is asked to

carry out at the station. At the Medical College of Pennsylvania, we have found 10 minutes stations quite appropriate to test basic skills of our junior medical students. Time constraints posed by the total duration of the examination must be carefully considered, otherwise lengthy stations will lead to limited sampling of the curricular content and clinical skills. It is also advisable to allocate the same time for each station, thus facilitating the smooth movement of examinees from one station to the next. Transit time must be built into the total time allocation for each station, e.g. for a 10 minute station, 9 minutes may be allocated for the task given to the examinee and 1 minute for the examinee to move to the next station. This additional time also allows examiners to complete their checklists prior to entry of the next examinee. Feedback stations may be included to provide examinees with immediate feedback on performance at the previous station. "Rest stations" may be incorporated between distant examination stations, to allow smooth transition from one station to the next. It is also prudent to prepare an extra station for each examination which can be set up with minimal effort at the last minute in case one of the regular stations fails to materialize, e.g. a real patient may not be available, a standardized patient may be sick, etc. This additional reserve station may incorporate the use of a plastic model or laboratory/x-ray studies.

Each examination station requires a number of details to be addressed. The skill to be assessed at that station will determine whether one requires a real patient, a standardized patient, laboratory data, roentgenographic studies or other props. Real patients with acute problems do not lend themselves well to this type of examination; however, those with chronic stable physical findings (funduscopic changes, goiters, adventitious pulmonary sounds, cardiac murmurs, abdominal organomegaly, ostomies, skin changes, deformities) can be used very effectively. They may also be used in history-taking stations. Real patients are best obtained through the assistance and cooperation of their primary physicians. Standardized patients have several distinct advantages over real patients. They can be carefully controlled and their history is more reproducible (2). Also, their simulation of physical findings can be standardized. These patients are especially useful in the evaluation of skills where participation of real patients would cause them undue distress or embarrassment. Using standardized patients, the level of difficulty of the examination station can be more readily controlled and their use during subsequent examinations can allow accurate comparison between the standards of performance of different students (1). Standardized patients can also be trained to evaluate firsthand and provide feedback on such issues as display of professionalism, respect, and gentleness of technique.

Complete instructions must be written for the examinee, stating clearly what the examinee is expected to carry out at that station. These must provide all relevant information in a concise, unambiguous fashion. Instructions must also be written for the patient (real or standardized) conveying to the patient what he/she is required to do and how much information is to be given to the examinee during the course of the examination. Another set of instructions must also be prepared for the faculty members who will be evaluating student performance during the OSCE.

Performance of examinees is usually evaluated by faculty or trained standardized patients using a checklist. Such checklists must be prepared for each station where they will be used. Preparation of checklists requires predetermined objective criteria

that are agreed upon by the examination committee, based on faculty input. A typical checklist has the key items listed which the examinee is expected to perform at the station. Different columns allow a check mark to be placed in one column if the item was satisfactorily carried out and in the other column if it was attempted but not carried out satisfactorily. At the completion of the examination, the total number of checks in each column are added to determine the overall performance of the student at that station. Additional bonus points may be given for overall technique and the examinee's attitude towards the patient. These points should, however, be limited owing to the somewhat more subjective nature of such evaluations. Scoring of each checklist, must be carefully considered and agreed upon by the examination committee. Measurement issues relating to the OSCE have been covered elsewhere in this manual.

If the examinee is required to answer multiple choice questions or fill in appropriate responses on an answer sheet, these too must be prepared for the examination station. To avoid cueing of the examinees, ideally these questions or response sheets should be provided at a separate station following the station where the specified task has been carried out (2).

A list of items required for each station must be prepared. This should be comprehensive and include all medical equipment and stationary needs.

It is also the responsibility of the examination committee to select an appropriate site to conduct the OSCE. Some institutions have special teaching/evaluation facilities, specially designed and equipped for this purpose. If such a facility is not available, the examination may be conducted in the outpatient department or any other space where several offices are available in close proximity to each other. Some stations, where the examinee only has to answer written questions, may be set up in the hallway. While selecting a suitable site it must be borne in mind that the buzzer or bell used to indicate the time to switch stations must be clearly audible at all stations.

An OSCE requires considerable clerical support through all phases of implementation. It has been estimated that approximately 80 hours of secretarial time is required to organize a large scale OSCE for the first time in an institution. This clerical assistance is absolutely essential and must be provided to the examination committee.

Additional funds required to support such an examination must also be considered. Use of standardized patients can add substantial cost to this endeavor. In some places, Collaboration with local amateur actors and actresses has yielded good results, without the need for hiring the more expensive standardized patients for the same purpose.

The examination committee must meet regularly and should have all personnel Involved with the examination – examiners, patients, etc. – identified and committed four weeks prior to the scheduled OSCE. During the subsequent three weeks, all standardized patients should be trained and validated.

Detailed instructions relating to the examination should be sent to all examiners, examinees and patients. Also, all examination materials including checklists,

examinee response sheets and instruction sheets must be finalized. Additional equipment required for the examination must be obtained at this time. Approximately one week prior to the examination, all personnel participating in the examination should be sent reminders and equipment items, as well as examination materials, must be rechecked.

## Conducting an OSCE

About 1 hour prior to commencement of the examination, each station should be set up with the appropriate equipment and examination materials. Instructions defining the student task at each station are left outside the room or at an appropriate place inside. Each station is numbered and the buzzer or bell checked. Examiners, examinees and patients are asked to arrive at the examination site at least ½hour prior to the examination. Orientation sessions are given to all 3 groups separately. Prior dissemination of detailed information relating to this examination greatly facilitates this process.

When all examination materials are in place and patients as well as examiners ready, examinees are asked to go to their pre-assigned stations. They are told to wait in front of their respective stations until the sound of the buzzer or bell indicating they may begin the station. At each subsequent sound of the buzzer or bell, examinees move to the next station. The process stops only when students have rotated through all stations. If the first station assigned to an examinee is based on information that the examinee should have obtained at the previous station, the examinee does not start the examination until the next rotation of stations. This person will finish the examination at the first assigned station, a little later than the others. If a test station is followed by a feedback station, the examiner moves with the examinee to the feedback station and another examiner takes his/her place in the test station. When this latter examiner moves, in turn, to the feedback station, the former goes back to the test station and the cycle continues.

Patients, whether real or simulated, may be able to tolerate only a certain number of physical examinations. Hence, at appropriate intervals, the clerical assistants may be required to relieve them by bringing in other patients with similar findings. This change can be accomplished smoothly during the time examinees are moving from one station to the next. It is also the responsibility of these assistants to attend to other needs of patients while the examination is in process.

Checklists, as well as student answer sheets, are collected by the clerical assistants either during the course of the examination or at the completion of the examination.

## OSCE Results

Results of the examination are based on previously determined criteria. If performance of an examinee is graded by more than one examiner at a station, the grades of the examiners should be averaged. Generally, good correlation has been found between these sets of grades. Computers may be used to obtain results in an

expeditious fashion. Thus detailed feedback may be provided to examinees soon after the examination. Generally students have favored receiving immediate feedback. Experience has shown that this can be provided during the course of the examination without adversely affecting the performance at subsequent stations.

At the Medical College of Pennsylvania, we have sent the results of student performance at each station back to the faculty member who designed that station. Likewise, dissemination of this information to appropriate course directors can provide invaluable information relating to the effectiveness of different courses. The OSCE can serve as a very powerful tool to drive the curriculum.

### Advantages of OSCE

The OSCE format allows a wide variety of clinical skills to be within a relatively short period of time (1,3,4). The content and complexity of individual stations can be specifically tailored to meet the educational needs of a course or an institution. All examinees are exposed to similar clinical material and valid comparison between the performances of different examinees can be drawn both efficiently and accurately. The objective and standardized nature of this examination helps to minimize examiner bias, resulting in greater inter-examiner reliability (3). Thus more junior or relatively inexperienced examiners may be used, if required, without jeopardizing the reliability of evaluations. (1). As examinees are exposed to a number of different stations, with different Examiners, each station is in effect a separate “mini-examination”. Consequently, poor performance at one station results in less stress for the examinee, who knows that the negative effect on overall performance may be minimized by performing better at other stations. Immediate feedback may be provided during the OSCE (5). Thus, OSCE can function as a powerful learning tool.

OSCE has been found to be a valid, reliable and practical method to assess clinical competence (1,3,4,6). It has been generally well received by both examiners as well as examinees (3,4,7).

### Disadvantages/Limitations of OSCE

As components of clinical competence are broken down for examination at different stations, the OSCE format leads to their assessment in a compartmentalized fashion. There is no opportunity to observe the examinee carrying out a complete evaluation of the patient (1,8). If stations are too short, the problem may get compounded resulting in difficulty with assessment of even the more simple clinical skills. It may also be difficult to assess certain skills, since repeated patient examinations may result in too much discomfort to the patient. Emergency situations do not lend themselves to this format of skills assessment and examinee attitudes are very difficult to assess during an OSCE. One might get some idea about examinee attitudes during his/her performance at a station, but generally these are better evaluated over a prolonged period of close observation (9). Although the OSCE is

generally accepted as a valid and reliable method of evaluation, the overall validity and reliability of the examination is totally dependent upon the quality of individual stations.

The setting of OSCE is somewhat “artificial”. Hence, faculty, as well as students, need to be acquainted with this format for better acceptance of the methodology.

An adequate pool of stations is required for repeated examinations. This may be limited owing to financial constraints or other logistical difficulties.

Although some examinees have expressed greater pre-examination, as well as Intra-examination, tension during an OSCE as compared with other forms of Examination (8), familiarity with this format decreases examinee anxiety. We have observed the students more at ease after rotating through the first few examination stations. An improvement in the scores of examinees has also been observed the course of rotations through OSCE stations, possible as a result of decreasing anxiety or increasing familiarity with the examination format (10). The OSCE can be quite fatiguing for the patients. Hence, patients should be selected carefully and several patients may be required for a single station if there is a large number of examinees or the examination particularly taxing (1). Examiners may also find this to be a grueling exercise, but the more efficient use of examiners’ time adequately compensates for this disadvantage (3).

In addition, compared with other traditional methods of evaluation, OSCE requires more preparation time and considerable clerical support (1,3,4).

### **Applications of OSCE**

This examination format has been used both for in-training evaluations, as well as final examinations by many medical and surgical disciplines. It may be utilized to assess clinical competence of undergraduate or postgraduate students. In-training evaluation using the OSCE can specifically identify areas of weak performance, facilitating remedial action to correct deficiencies.

OSCE may provide valuable information relating to whether course objectives are being accomplished. Thus appropriate changes in the curriculum can be made to better address the needs of students. This approach has been effectively utilized at our medical school.

Clinical skills of students from different institutions have been compared using OSCE stations (11) and these stations have been placed in databanks. Such OSCE databanks have been established to pool resources and allow easy access to available material. Ideally, OSCE should be used in conjunction with other methods of evaluation to provide comprehensive information on the clinical skills of examinees.

## References

1. Harden, R.M. and Gleeson, F.A. Assessment of Clinical Competence Using an Objective Structured Clinical Examination (OSCE). Medical Education 13: 41-54, 1979.
2. Harden, R.M. Stevenson, M., Downie, W.W., and Wilson, G.M. Assessment of Clinical Competence Using Objective Structured Examination. British Medical Journal 1:447-451, 1975.
3. Cuschieri, A., Gleeson, F.A., Harden, R.M. and Wood, R.A.B. A New Approach to a Final Examination in Surgery. Annals of Royal College of Surgeons of England 61:400-405, 1979.
4. Watson, A.R., Houston, I.B. and Close, G.C. Evaluation of an Objective Structured Clinical Examination. Archives of Disease in Childhood 57: 390-398, 1982.
5. Back, N.M.I. and Harden, R.M. Providing Feedback to Students on Clinical Skills by Using the Objective Structured Clinical Examination. Medical Education 20:48-52, 1986.
6. Hart, I.R., Cairncross, R.G., Harden, R.M. An Objective Structured Examination of Clinical Competence. Annals of Royal College of Physicians and Surgeons of Canada 12: 58, 1979.
7. Lazarus, J. and Kent, A.P. Student Attitudes Towards the Objective Structured Clinical Examination (OSCE) and Conventional Methods of Assessment. South African Medical Journal 64 (11): 390-394, 1983.
8. Kirby, R.L. and Curry L. Introduction of an Objective Structured Clinical Examination (OSCE) to an Undergraduate Clinical Skills Programme. Medical Education 16: 362-364, 1982.
9. Sachdeva, A.K. Objective Structured Clinical Examinations (OSCE). In: Merrick, H.W. (ed.), Manual of Evaluation Techniques for Medical Students. Springfield: Association for Surgical Education, 91-103, 1987.
10. Ross, J.R., Syal, S., Hutcheon, M.A., Cohen, R. Second-Year Students' Score Improvement During an Objective Structured Clinical Examination. Journal of Medical Education 62 (10) 857-858, 1987.
11. Hart, I.R., Cairncross, R.G. and Harden, R.M. The Use of the Objective Structured Clinical Examination to Compare Clinical Skills in Two Medical Schools. Medical Education 15: 64, 1981.

## HOW TO PREPARE AND IMPLEMENT AN OBJECTIVE STRUCTURED CLINICAL EXAMINATION

R. Cohen, Ph.D., A MacInnes, B.Sc.,  
J. Finkelstein, RNNP, A. I. Rothman Ed. D.,  
A. Robb, BA

### Introduction

*Concern about the quality of the assessment of clinical skills has given rise to intensive research into the development of methodologies to better evaluate these skills. Harden et al, 1979, are credited with the introduction of the Objective Structured Clinical Examination (OSCE) as an evaluation method which is viewed as having greater potential for producing both reliable and valid assessments of these skills.*

*The OSCE is a testing format that facilitates the assessment of relatively large number of students on numerous clinical cases. Standardized clinical situations are organized into a circuit of stations, half of which require the student to interview or physically examine patients, and half requiring the student to answer questions about the patients that were seen or similar cases. The time at each station is limited, thus allowing a large number of students to be assessed in a relatively short period of time. Students at each of the patient-based stations are observed by physicians who make use of pre-formulated checklists to evaluate either the history-taking or physical examination skills.*

*Since its introduction, there has been increasing interest in the OSCE by medical educators in both North American and other parts of the world. The OSCE has been used to evaluation the clinical skills of medical students in Introduction to Clinical Methods courses, in systems teaching, interviewing and communication courses, in clinical clerkships and in residency programs as both formative and summative instruments.*

*While much more research is required to assess its reliability and validity, there is little doubt that the OSCE provides medical educators with a more reliable source of information concerning the clinical performance of both undergraduate medical students and residents. It has been shown to be superior to the more traditional methods of global assessments and long case oral examinations. In addition, the OSCE provides both student and faculty with in-depth feedback. For the student, the almost instantaneous feedback provides the information necessary for remedial work when the OSCE is used as a formative assessment tool. For the faculty, it provides insights into the effectiveness of the teaching of clinical skills and identifies areas of omission in the curriculum. The original OSCE, introduced by Harden et al, was based upon the use of multiple short (3 to 5 minute) stations. Subsequent developments in the OSCE have led to diversity, both in terms of the duration of the station, and in terms of skills that can be assessed at one station. Most recently, medical educators at Southern Illinois University, University of Manitoba and the University of Massachusetts have implemented the Multiple Long Station*

*examination. This variation on the OSCE theme has students spending up to 45 minutes at a station. This type of examination is viewed as being more representative of the interaction that takes place between doctor and patient than is the short (5-minute) station. The question of which of these is the better testing format has yet to be answered.*

*The OSCE which is described in this chapter is based on the experience of the Clinical Skills Assessment Unit of the Faculty of Medicine at the University of Toronto. This Unit is responsible for the development and implementation of an evaluation program for the selection of Foreign Medical Graduates for a government-sponsored licensing program.*

### **An Example: The 1987 CSAU-OSCE**

The OSCE consisted of both 5 and 10-minute stations as well as 10-minute rest stations. The 5-minute stations were arranged in twenty-one couplets (pairs). In each couplet, history or physical examination skills were assessed in the first part, data interpretation and problem-solving skills in the second part. The nine 10-minute problem-solving stations required the candidates to respond to short open-ended questions relating to written scenarios and relevant artifacts (X-rays, 35-mm slides, and photographs). In Table 1, the clinical situations used in the OSCE, as well as the more general skill areas. (History-taking, physical examination, problem-solving, interviewing, etc.), are listed.

Standardized patients were used in those stations where history-taking, physical examination interviewing skills were assessed. At least two patients were provided for each station. The patients were rotated systematically according to a preset schedule. History-taking, physical examination and interviewing skills were evaluated by physicians utilizing checklists. Each station was staffed by at least two physicians, although only one physician at a time served as an evaluator. Previous research has demonstrated that inter-rater agreement between qualified examiners is high (2)

The seventy-two candidates were randomly assigned to two groups of equal size that were examined separately in the two day Clinical Skills Examination. Group one did Part One of the examination in the morning of the first day, and Part Two in the morning of the second day. Total examination time per candidate, excluding rest stations, was five hours (2 1/2 hours per day). The twenty-one couplets and nine 10-minute stations provided thirty 10-minute units; each scaled to provide a twenty point maximum score. The thirty scores were summed to provide a total score with a maximum possible value of 600.

## How to Develop an OSCE

The decision to run an OSCE generally begins when one or two faculty members decide that they want to pilot test the use of the OSCE as a means of assessing the clinical skills of students. If the OSCE is developed and implemented by one person, it will in all probability not be attempted again. The process of developing an OSCE is one that is best accomplished by a team.

### The Coordinating Committee

Ideally, this Committee is made up of a coordinator (either M.D. or educator), a small number of physicians who will serve as the key decision makers, a larger group of physicians who will be given responsibility for the writing of stations, and a patient coordinator/trainer who will assume the responsibility for the training of both real and simulated patients. Secretarial and clerical assistance is essential for typing and reproducing the material to be used in the examination. The presence of an educator who is skilled in the area of testing and who is able to contribute to the decisions that need to be made relating to the issues of reliability and validity is important.

### The First Meeting

One of the first issues that needs to be dealt with is the objective of the examination. The purpose of the examination plays a role in determining the number of stations that makeup the examination. If the examination is a pilot test, or if it is to be used as a diagnostic tool rather than a final assessment of students clinical competence, the number of stations is less of an issue. If, however, the examination is to be used for a comprehensive summative evaluation, the number of stations is important. Research that has been carried out on the OSCE format reveals that in order to achieve an acceptable level of reliability ( $>0.80$ ) for the assessment of history and physical examination skills, it is necessary to have at least a day to 1 1/2 days of testing time (3,4).

Once the number of stations has been decided, the Committee needs to select the skills that are to be evaluated, the topic areas that are to be covered, and the level of difficulty of each station in the examination. This factor plays an important role in determining the overall reliability of the examination. Stations that are too easy will provide an examination with an unacceptable level of reliability. The number of history and physical examination stations must be set. For example, if the examination is to test surgical content only, the distribution of stations by subspecialty must be determined. Harden et al designated 50% of the stations to history-taking and physical examinations skills, and the remainder to testing problem-solving skills related to the patient problem seen.

Finally, if the examination is to be used to determine whether a student passes or fails, criteria and standards must be established. It may, for example, be decided that the student must satisfactorily complete a certain percentage of the stations in order to pass the examination. The student may be expected to achieve a certain minimum

mark on each of the stations in order to pass. It is also possible to determine what remedial action will be demanded of a student who fails to meet the prescribed standards.

In the case of the CSAU 1987 OSCE, the skills tested were as follows:

Data gathering skills

Data interpretation skills (X-rays, lab results, etc.)

Ability to formulate investigation and management plans

Problem-solving skills

Interpersonal skills

Due to the fact that selection decisions had to be made in this examination, the number of stations had to be sufficient to provide a level of reliability  $>0.80$ . The decision was made to have 21 couplets (42x5-minute stations) and nine 10-minute stations. Based on this number of stations, each candidate spent 2 1/2 hours of testing time on each of two testing days. Once the decisions had been made by the Coordinating Committee relating to the number of stations, the content of the stations, and the skills to be tested at each station, the group of physicians who were to be given the responsibility of authoring the stations was convened.

## **The Authoring Team**

The Authoring Team consists of physicians from all of the specialties that are to be included in the examination, the educator patient coordinator, and an administrative assistant. Members of the Authoring Team should be provided with a detailed explanation of the OSCE format and relevant readings prior to the meeting. At an orientation meeting, the Team should be presented with a discussion and clarification of the OSCE format. The purpose of the examination as envisaged by the Coordinating Committee should be discussed and the rationale for the distribution of topics and skills should be explained.

In all probability, the physicians called upon to participate in the authoring of questions will have had minimal experience in writing OSCE stations. It is essential that these physicians be given adequate guidelines for authoring a station.

It is also essential to dispel the idea that the station can be written in one short sitting. There are a number of ways of assisting the author in their task. The most simple way is to provide them with copies of stations that have been used effectively in other OSCEs. Ian Hart at the University of Ottawa has developed a bank of OSCE stations covering many specialties. This bank of stations can be purchased, and if necessary, may be used freely. However, the exercise of creating stations locally has benefits that outweigh the intensive time requirements necessary for this process.

## **Guide for Creating a Station**

A couplet station (2x5-minute stations) consists of the following sections: (Note-Appendix No. 1 contains examples of a history-taking station, physical examination station, 10-minute interviewing station and a 10-minute problem-solving station.)

A sheet of instructions to the candidate.

A checklist for the assessment of the history/physical examination with a scoring/weighting scheme.

A set of questions to test the problem-solving skills of the candidate. These should not be M.C.Q questions.

Answers to the questions asked and a scoring scheme.

A detailed patient profile for the training of a standardized patient, if such a patient is to be used.

A listing of the equipment and artifacts (X-rays, lab result, and instruments) that will be required at the station.

### **The Instructions to the Candidate**

This sheet contains a brief statement of the problem that the student will encounter at the station. It should be stated in clear and unambiguous terms and should not be too long. The student will have no longer than one to one and a half minutes between to read the instructions and prepare him/herself for the demands of the station.

### **The History/Physical Examination Checklist (The First Five Minutes)**

The checklist will be used by the observing physician to evaluate the history-taking, physical examination or interviewing skills of the students in those station where patients are utilized. The manner in which the checklist is written can contribute to the reliability of the station. To some degree, the more detailed the checklist the greater the power of the station to differentiate effectively among the abilities of students. Each item on the checklist should be followed by a suggested score. The score given to each item should reflect its' relative important. Each station can have its own total score which can be scaled appropriately when the results are entered for analysis. It is also appropriate to leave the evaluator to comment regarding the student's skills.

## **The Problem-Solving Questions ( The Second Five Minutes)**

The second part of the couplet is devoted to assessing the ability to interpret information. The information provided here could be any of the following: X-rays, C.T. scan, I.V.P. results, laboratory results, slides, or any other type of medical test or specimen. The questions asked in this section should not be of the MCQ type, but rather short-answer type questions. (See Appendix 1). The questions in this section may also deal with investigations, management plan, or differential diagnosis.

To facilitate the marking of the problem-solving component, a complete answer for each question must be provided together with the scoring scheme. This is essential if the station is to be marked by someone other than the author of the station.

## **Instruction for the Standardized Patient**

Standardized patients can be either healthy people who are trained to simulate a particular disease entity or chronically ill patients with stable findings who are trained to present their problem in a consistent and reliable manner. Regardless of whether a healthy or chronically ill standardized patient is to be used at a station, a detailed patient profile needs to be written by the author of the station. The profile must be written so that it is understandable to the lay person who may be presenting the role. An important component of the instruction package is a fact sheet. This must contain all of the historical information which the patient may be required to provide. In the case of a physical examination station, the package should contain a description of the signs to be elicited on examination, and a listing of the responses to be given by the patient when he/she is being palpated or examined for movement, etc. The instruction package must also contain information documenting how the patient should dress and behave in the station.

The sequence of training session that the patient trainer must implement in the process of preparing patients for the OSCE is outlined in Appendix 2 (Standardized Patients: The Training Process). It is essential that the patient trainer meet with each of the station authors to discuss the patient profiles and thereby obtain a detailed understanding of the problem to be presented. It is also desirable for the author of the station to meet with the standardized patient during the training period. This is especially important when a patient is being trained to undergo a physical examination.

The number of hours required to train standardized patients is discussed in greater detail in Appendix 2. Our experience indicates that a standardized patient can be trained in two-hour sessions to present a history. For a physical examination (eg. examination of the knee), one session is required. A more complex physical examination would require an additional one or two sessions. The training of patients for a 10-minute interviewing station may require between 3-4 hours.

## **Equipment Listing for Stations**

The station authors must provide a detailed list of the artifacts (X-rays, C.T scans, laboratory results, photographs, etc.) that will be required for each station. The listing of equipment such as blood pressure cuff, stethoscope, etc., must also be provided.

## **Patients for the OSCE: Real or Simulated**

Each OSCE station that has as its objective the assessment of history-taking or physical examination skills will require the presence of a standardized patient. A standardized patient is either a healthy individual or a person with a real health problem that is characterized by stable chronic findings. In both cases, the individual must be trained to present their problem repeatedly and consistently (5).

In recent years, the validity of assessment procedures that have utilized healthy individuals as standardized patients has been well documented (3,6,7). It has also been shown that a well-trained standardized patient can be used not only to present the patient problem in a consistent and reliable manner, but also to evaluate the clinical skills of the student. Standardized patients have been used to evaluate student clinical skills in OSCEs in a number of medical schools (3,4).

## **Follow-up Meeting**

A critical path for the development of an OSCE examination should be drawn up. It is essential that deadlines be set and kept in the development stage. An important feature of the development phase is the maintenance of contact between the OSCE coordinator and the authoring team. Once the authors have completed their first draft, they should be typed in a uniform format. The draft OSCE station should then be reviewed by the OSCE coordinating team. The review is guided by the objective of the examination. All components of every station must be listed for each station author. This process should continue until the OSCE coordinating committee is satisfied that each station is representative of the objective set for the examination and meets the criteria set out for a 5 or 10 minute station.

## **The Training of the Examiners**

At the station where clinical skills (history-taking, physical examination, interviewing and communication) are assessed, the evaluation may be done by either a physician or a standardized patient. If a standardized patient is used, he/she is trained to present both the role and do the evaluation. In both cases, the use of a predetermined checklist is essential for the maintenance of reliability. Research has shown that an acceptable level of reliability can be achieved with either a physician or standardized patient as the examiner (3,4,8).

In the case of the CSAU 1987 OSCE, station authors were requested to provide two physician examiners for their stations. Only one physician at a time served as an evaluator. A number of weeks prior to the examination, the two examiners were provided with a complete description of their respective stations. A meeting between them and the station author was held during which each item on the checklist was reviewed and criteria for awarding the candidate the allocated marks were identified. It is essential that there be agreement between the examiners and the authors about the criteria to be used for assessing the candidates.

Ideally, a training session should be held for the two examiners together with the standardized patients for that station. The examiners together with the standardized patients for the station. The examiners should be given the opportunity to observe the history-taking or the physical examination to be conducted, to score the event independently, and then to compare their individual scores. This process provides the two examiners with the opportunity to achieve agreement on the criteria to be used. It also enables the standardized patients and the patient trainer to gain a better understanding of the expectations for the successful implementation of the role.

### **Choosing a Site for the OSCE**

The setting that has been found to be most suitable for the running of an OSCE is an outpatient department or a clinic. These settings generally consist of a large number of small offices and examination rooms that are ideal for the OSCE. They are generally laid out in a physical plan that places them in close proximity to one another and they are equipped with the instruments that are needed. It is important to keep in mind that students are required to move from station to station during the examination. If the rooms are not physically close, change-over periods will be hectic and their coordination will be difficult. It has been our experience that when an OSCE is planned for a large number of students, an outpatient department setting on a weekend provides the ideal environment.

### **Running the OSCE**

The number of support staff required to run the OSCE is dependent upon the number of stations in the examination. The personnel that are essential for the efficient running of the OSCE are:

1. An official timer, with two stop watches, a bell, or whistle.
2. One support person per three stations. The responsibility of these individuals is to collect completed answer sheets and generally see that the station is ready for the next candidate during the change-over period.
3. A clerical person at a central desk who is responsible for the sorting of the completed checklists and answer sheets.

4. A chaperone for each of the rest stations. It is their responsibility to accompany the candidates to the washrooms and to provide refreshments.
5. The patient trainer who should be circulating to deal with any patient problems.
6. At least two reserve standardized patients who have been trained to assume a number of roles.
7. A number of reserve stations that can be substituted in event of difficulties. These can be stations that do not require the use of patients.

### **Setting up the OSCE**

The examination setting must be prepared in advance. The following steps should be carried out in the weeks prior to the examination day:

1. A map of the clinic should be drawn. The map should identify the exact location of each office and examination room.
2. Stations should be allocated and marked on the map, so that stations can be easily identified and marked on the examination day.
3. The traffic flow of the student through the stations should be outlined on the map. This can be easily marked on the appropriate spots in the clinic on the examination day.
4. All material for identifying the stations should be prepared. Each station should be identified with a sign that gives the station number. Rest stations must be clearly marked.
5. All stations that require special equipment should be prepared and the equipment placed in the station.
6. An examination booklet should be prepared for each candidate. This booklet should contain the Instruction to Candidate Sheets for each station. On the front sheet, the student name and the number of the station at which he/she will start the examination is recorded. This allows the student to read the appropriate instruction to Candidate Sheet during the change-over period in preparation for the upcoming station. This sheet briefly identifies the task to be carried out at the station. The student may also record information obtained from the patient for use in the second part of the couplet.
7. Contingency plans need to be drawn up for situations where a student is required to leave the examination. Examiners, helpers and standardized patients should all be aware of how the examination will progress. Once the bell has sounded the start of the OSCE, it should run until its completion without interruption.

8. Students should be randomly allocated to the stations they will begin at the start of the examination. If two groups of students are to be rotated, they should be randomly assigned to groups.

**REFERENCES**

1. Harden, R.M. and Gleeson, F.A. Assessment of Clinical competence Using an Objective Structured Clinical Examination (OSCE). Medical Education 13:41-54, 1979
2. Harden, R.M., Stevenson, M., Downie, W.W. and Wilson, G.M. Assessment of Clinical competence Using Objective Structured Examination. British Medical Journal 1:447-451, 1975.
3. Stillman, P.L., Swanson, D.B., Smee, S., et al. Assessing Clinical Skills of Residents with Standardized Patients. Annals of Internal Medicine 105:762-771, 1986.
4. Petrusa, E.R., Blackwell, A., Parcel, S. and Saydjari, C. Psychometric Properties of the Objective Clinical Examination (OSCE) as an Instrument for Final Examination, in Hart, I.R., Harden, R.M. and Walton, H.J> (eds.) Newer Developments in Assessing Clinical Competence. Montreal: Heal Publications 1986.
5. Barrows, H.S. Simulated Patients. Springfield: Charles C. Thomas, 1972.
6. Stillman, P.L., Rugill, J.S., Rutala, P.J. and Sabers, D.L. patient Instructors as Teachers and Evaluators. Journal of Medical Education 55:186-193, 1980
7. Norman, G.R. and Tugwell, P.A. Comparison of Resident performance on Real and Simulated Patients. Journal of Medicine Education 57:708-715, 1982.
8. Robb, K.V., Rothman, A.I. The Assessment of Clinical Skills in General Medical Residents: Comparison of the Objective Structured Clinical Examination with a conventional Oral Examination. Annals of the Royal College of Physicians and Surgeons in Canada 18:235-238, 1985.
9. Petrusa, E.R. et al, "An Objective Measure of Clinical performance". The American Journal of Medicine 83:34-42, 1987.
10. Newble, D.I. and Elmslie, R.G. A new Approach to the Final Examinations in Medicine and Surgery. Lancet 8245: 517-518, 1981.
11. Newble, D.I., Hoare, J. and Elmslie, R.G. The Validity and Reliability of a New Examination of the Clinical Competence of Medical Students. Medical Education 15:46-52, 1981.
12. Harden, R.M. and Cairncross, R.G. Assessment of Practical Skills:
13. The Objective Structured Practical Examination (OSPE). Studies in Higher Education 5 (2): 187-196, 1980
14. Linquist E. F. (ed.), Educational Measurement. Washington, D.C.: American Council on Education, 1951.

**Table 1**

The clinical Skills Examination: Clinical Situations and Unit Types

<u>Situation/Couplet</u>	<u>Type</u>
1. Well Baby Care	H/Ps
2. Vaginal bleed	H/Ps
3. Breast Examination	P/Ps
4. Level of Consciousness	P/Ps
5. Shortness of Breath	P/Ps
6. Shortness of Breath	H/Ps
7. Diarrhea	H/Ps
8. Intermittent Claudication	P/Ps
9. Extracellular Fluid volume	H-P/Ps
10. Back Injury	P/Ps
11. Left Lower Quadrant Pain	P/Ps
12. Pregnancy	H/Ps
13. heart Murmur (pediatric)	S/Ps
14. Anemia	H/Ps
15. Splenomegaly	P/Ps
16. Knee Examination	P/Ps
17. Chest Pain	H/Ps
18. Left Flank Pain	H/Ps
19. Spinal Cord Injury	P/Ps
20. Thyroid	P/Ps
21. Back Pain	Lab/Ps

<u>Situation/10 Minute Units</u>	<u>Type</u>
22. Lightheadedness	Ps
23. Anorexic Adolescent	I
24. Post-operative Fever	H/P
25. Contraception	I
26. Acute Respiratory Distress	Ps
27. Bad News.	I
28. Shortness of Breath	Ps
29. Oncology (Hypercalcemia)	Ps
30. Mental Status	H

Couplet units - units 1 to 21; 10 minute stations - units 22-30.

Day one - units 1 to 14; Day two - units 15 to 30

Unit types: H - history, P - physical examination, I - interview,  
Ps - problem-solving, S - mechanical simulation,  
Lab - laboratory procedure.

## **APPENDIX 1**

Station 15 - Day 1

Instructions to Candidate

This woman presents to your office with a complaint of cough and shortness of breath.

OBTAIN A **FOCUSED** HISTORY.

At the next station you will be asked to answer some questions from information you have acquired from this patient or from information that will be provided.

**APPENDIX 1**

## Station 15

## History Scoring Sheet

ITEMS	Max Score	Check if Obtained
Age	1	_____
Onset of symptoms	2	_____
Duration of cough	2	_____
Day vs. night cough	2	_____
Cough productive/non-productive	2	_____
Wheezing	2	_____
Aggravating factors - exercise	1	_____
cold	1	_____
noxious fumes	1	_____
Environment	1	_____
Chest pain	1	_____
Duration of s.o.b.	2	_____
Orthopnea	2	_____
Deep breathing	1	_____
Hemoptysis	2	_____
Exercise tolerance	2	_____
Swelling of ankles	2	_____
Paroxysmal nocturnal dyspnea	2	_____
Allergies - past	1	_____
- family	1	_____
Pets/Birds	1	_____

Smoking history	2	_____
Arthritis	2	_____
Skin lesions	1	_____
Dysphagia	1	_____
Use of drugs	2	_____
Acts courteously and respectfully to the patient	1	

**Maximum Total Score (41)**

Examiner\_\_\_\_\_

Patient\_\_\_\_\_

Total Score\_\_\_\_\_

**APPENDIX 1**

Station 16

Answer sheet

1. Review the chest x-ray provided and list two abnormal findings.
2. Interpret the pulmonary function tests. (Next page)
3. List three possible diagnoses in order of likelihood.

Marker\_\_\_\_\_

Total Score\_\_\_\_\_

## APPENDIX 1

### PULMONARY FUNCTION TESTS

#### Station 16

POSITION DURING TESTS	Predicted	Observed	Percentage	Post-Bronchodil.	
Erect ( ) Normal	Supine ( )	Normal	Value	of normal	Observed %
Forced Vital Capacity	L	5.20	3.41	66	
Forced Expiratory Volume 1 - second	L	4.03	2.81	70	
Ratio of FEV1/FVC	%	78	83		
Lung Volumes:					
Functional Residual Capacity	L	3.51	2.50	71	
Residual Volume	L	2.03	1.43	71	
Total Lung Capacity	L	7.01	5.30	76	
Ratio of RV/TLC	%				
CARBON-MONOXIDE (SS)					
DIFFUSING CAPACITY	<u>ml/min</u> mmttg	32.71	17.71	54	

**APPENDIX 1**

## ANSWER SHEET

## Station 16

1. Review the chest x-ray provided and list two abnormal findings.

Interstitial lung disease 5

Bilateral hilar lymphadenopathy 5

2. Interpret the Pulmonary Function tests.

Normal expiratory flows 1

Decreased lung volumes 1

Low DLCO 1

Restrictive disease 1

3. List three possible diagnoses in order of likelihood.

Sarcoidosis 4

Lymphoma 3

Lymphangitic carcinomatosis 2

Collagen vascular disease, Wagner's 1

External allergic alveolitis 1

Idiopathic pulmonary fibrosis 1

Maximum Total Score (28)

If sarcoidosis is the #1 diagnosis - score 2 additional marks.

If lymphoma or lymphangitic carcinomatosis is #1 diagnosis -  
score 1 additional mark.

If others are #1 diagnosis - no additional marks.

## APPENDIX 1

### Station 16 – Respiriology

#### Instructions to Simulated Patient

Mrs. Jones, who is approximately 60 years old, presents to your office with a 4-month history of gradually progressive shortness of breath associated with a non-productive cough. Her present problem of cough and shortness of breath seemed to follow a 4-day flu-like illness approximately four months ago. The cold seemed "settle" in her chest afterwards.

She had worked as a saleswoman at Eatons for 30 years, but 3 months ago she had to take an early retirement because of her poor health. She just couldn't keep up the pace any longer.

She has no unusual hobbies, occupational exposure or pets including, budgies (she has never had any pets in her house). She has never smoked.

Her past health has been excellent. Hospitalizations have been for childbirth only. She has two married daughters. She has never had any surgery and no history of chronic illness.

No childhood history of allergies, eczema or skin rashes. She has no symptoms to suggest and seasonal or perennial rhinitis (running nose) or post nasal drip. She denies any history of hives or drug allergies. No history of arthritis, arthralgias, skin rashes or Rheynaud's phenomenon.

The cough is primarily a problem during the daytime and seems to be precipitated by breathing or exercise. Exposure to strong fumes such as care exhausts, cleaning agents and second hand smoke also bring on the cough. It is a dry cough, so there has been no sputum and also no hemoptysis (coughing up blood). The cough is also troublesome when she first lies down but, once she falls asleep, she is able to sleep through the night without any problems. She sleeps on one pillow and denies any orthopnea (difficulty breathing when lying flat), paroxysmal nocturnal dyspnea (shortness of breath during sleep). She denies any swelling of ankles, wheezing or chest pain.

One year ago she was able to play eighteen holes of golf, albeit with a golf cart, but walking at least half the holes. She was also playing doubles tennis without any problems.

At the present time, she is only able to walk approximately two blocks on level ground, without being significantly dyspneic, at a pace which holds up her husband. She has considerable problems with housework which was formerly not a problem. Activities requiring her to bend over are now very difficult and have to be done by her husband.

**APPENDIX 1****Station 7****Instructions to Candidate**

This woman is concerned about a possible lump in her left breast.

**CONDUCT A SYSTEMATIC PHYSICAL EXAMINATION OF BOTH BREASTS. DESCRIBE TO THE EXAMINER WHAT YOU ARE DOING AND YOUR FINDINGS.**

At the next station you will be asked questions related to the patient's problem.

## APPENDIX 1

Station 7

Breast Lump

ITEM	Scoring Key		Comments
	Done Correctly	Done Incorrectly	
1. INSPECTION			
Patient sitting, arms at slides	1	0	
Patient sitting, arms above head	1	0	
Patient leaning forward	1	0	
If only one breast examined		-1	
2. PALPATION			
Patient supine, arms behind head	1	0	
Systemic palpation of the four quadrants and tail	3	0	
Palpation of nipple, areola	1	0	
Palpation of axillary nodes	2	0	
Palpation of supraclavicular fossa	1	0	
If only one breast examined		-1	

Maximum Total Score (11)

Examiner\_\_\_\_\_

Patient\_\_\_\_\_

Total Mark\_\_\_\_\_

**APPENDIX 1**

## Station 8

## Instructions to Candidate:

A 42-year-old premenopausal woman presents to your office with a tender mass in her left breast. It is not associated with any other symptoms or signs. She has had "cysts" in the breasts in the past but is now concerned because her Mother (age 62) has recently had a mastectomy for carcinoma.

Your examination shows multiple thickenings in both breasts with a discrete, smooth, well-defined in the upper outer quadrant of the left breast. There are no other findings.

**QUESTIONS**

1. What is your differential diagnosis? Which is most likely?
2. What action should a specialist take at this time?
3. Review the mammogram. Describe the abnormality seen, and state the most likely diagnosis.

Marker\_\_\_\_\_

Total Score\_\_\_\_\_

## APPENDIX I

### Station 8

#### Scoring Key

A 42-year-old premenopausal woman presents to your office with a tender mass in her left breast. It is not associated with any other symptoms or signs. She has had "cysts" in the breasts in the past but is now concerned because her Mother (age 62) has recently had a mastectomy for carcinoma.

Your examination shows multiple thickenings in both breasts with a discrete, smooth, well-defined nodule in the upper outer quadrant of the left breast. There are no other findings.

#### QUESTIONS

1. What is your differential diagnosis? Which is most likely?  
FIBROCYSTIC DEASE (1)  
CARCINOMA (1)  
FIBROCYSTIC DIEASE MOST LIKELY (1)
2. What action should a specialist take at this time?  
ASPIRATION (1)  
MAMMOGRAPHY OF BOTH BREASTS (1)  
BIOPSY IF ASPIRATION UNSUCCESSFUL (1)
3. Review the mammogram. Describe the abnormality seen, and state the most likely diagnosis.  
IRREGULAR 2 CM MASS IN BREAST (1)  
CARCINOMA (1)  
Maximum Total Score (8)

**APPENDIX 1**

## INSTRUCTIONS TO CANDIDATE

## STATION B1

This is a problem-solving station.

Your task will be outlined in the information and/or materials provided.

**YOU HAVE 10 MINUTES TO READ THE INFORMATION AND ANSWER THE QUESTIONS. PLEASE ANSWER EACH QUESTION ON THE PAGE PROVIDED.**

**WRITE CLEARLY.**

## APPENDIX 1

### STATION B1

#### NEUROSCIENCE PROBLEM-SOLVING QUESTIONS

A 34-year old right-handed lawyer comes to your office complaining of a five-week history of lightheadedness and unsteadiness on her feet. Over the last ten days, she has had mild incoordination of the left side of the body. Yesterday her left side became noticeably weaker and more clumsy. She has been able to continue working at her job without any significant loss of efficiency. She is able to compensate for the left-sided clumsiness most of the time, tending only to scuff the left foot when she walks rapidly. She has some difficulty in the exercise classes that she attends three times per week. She denies headache.

On examination, she appears in good general health. You confirm that the abnormal physical findings are limited to the central nervous system. She is alert, oriented and cooperative. The visual fields are full to confrontation testing. There is no inattention to bilateral visual stimuli. The optic fundi are normal and good venous pulsations are seen. The pupils are equal and react to light. The external ocular movements are full and conjugate and there is no nystagmus.

With the arms in the outstretched posture, there is a mild drift of the left hand. The left grip is slightly weaker than the right. Strength in the legs is normal. The tendon reflexes in left are slightly accelerated with respect to the others which are normal. The plantar responses are flexor. On walking briskly in the hall she tends to scuff the left foot and does not swing the left arm as much as the right.

The sensory examination to light touch and pin prick sensation is normal. There is slight impairment of position sense in the left hand and she is intermittently inattentive to the left-sided stimulus on bilateral tactile sensory testing.

Her CT scan is provided.

ANSWER THE FOUR QUESTIONS THAT FOLLOW IN THE SPACE PROVIDED.

PLEASE REMEMBER TO WRITE CLEARLY AND TO PLACE YOUR NAME AND I.D. NUMBER ON EACH PAGE.

1. List 5 possible diagnoses.
  
2. What is the most likely lesion?
  
3. Why is this lesion your first choice?
  
4. Give 2 management options open to a specialists.

Maker\_\_\_\_\_

Total Score\_\_\_\_\_

**APPENDIX 1**

STATION B4

INSTRUCTIONS TO CANDIDATE

This 20-year-old woman comes to office requesting information regarding contraception.

TAKE A RELEVANT HISTORY AND ADVISE THE PATIENT ON CONTRACEPTION OPTIONS.

**APPENDIX 1**

## Station B4

This 20-year-old woman comes to your office requesting information regarding contraception. TAKE A RELEVANT HISTORY AND ADVISE THE PATIENT ON CONTRACEPTION OPTIONS.

THIS RATING SCALE TO BE USED FOR CATEGORIES A TO H:

- 3 marks - carried out all or most of the item under this heading in an adequate manner
- 2 marks - carried out at least half of the item under this heading in an adequate manner.
- 1 mark - carried out less than half of the items under this heading in an adequate manner.
- 1 ,marks - did not carry out any of the item and/or behaved in the incompetent and inefficient manner.

A. Initiation of the interview

- introduces self to patient 3\_\_\_\_\_
- Uses name of patient in greeting 2\_\_\_\_\_
- presents in a positive and attentive manner , establishes rapport 1\_\_\_\_\_
- attends to patient's comfort 0\_\_\_\_\_
- defines the purpose of the interview
- is at ease with the patient
- places patient at ease

B. Listening Skills Non/Verbal Communication

- maintains good eye contact
- has appropriate body posture 3\_\_\_\_\_
- makes empathic gestures 2\_\_\_\_\_
- uses silence appropriately 1\_\_\_\_\_

- is attentive to patient's agenda 0\_\_\_\_\_
- avoids distracting activities

C. Attitude

- avoids condescending, sychophantic or rude behavior 3\_\_\_\_\_
- appropriately congenial and non-judgmental 1\_\_\_\_\_ 0\_\_\_\_\_

D. Questioning Skills

- speaks clearly and fluently
- uses language understandable to patient 3\_\_\_\_\_
- asks open-ended questions 2\_\_\_\_\_
- asks one question at a time 1\_\_\_\_\_
- asks appropriate probing questions 0\_\_\_\_\_
- uses facilitating techniques
- controls the pace of the interview

E. Organization

- controls the direction of the interview 3\_\_\_\_\_
- logical flow of questions 2\_\_\_\_\_
- does not discuss management until history obtained 1\_\_\_\_\_ 0\_\_\_\_\_

F. Management Style

- does not make incorrect statements
- determines patient's level of knowledge and educates appropriately 3\_\_\_\_\_ 2\_\_\_\_\_ 1\_\_\_\_\_
- appropriate instructions are given in a readily understood manner 0\_\_\_\_\_

G. Management Content

- summarizes risks of oral contraceptive, IUD and barrier methods accurately and

### Succinctly

- recommends oral contraceptives and discusses common and important side effects, breakthrough bleeding, missed pills, drug interactions 3\_\_\_\_  
2\_\_\_\_  
1\_\_\_\_  
0\_\_\_\_
- recommends using condoms until protected by oral contraceptive
- recommends quitting smoking
- should follow-up patient with complete physical and pap smear, as well as checking for compliance with, side effects of medication

### H. Closing of Interview

- summarizes what has been said
- formulates a problem list (birth control, smoking, tension headaches)
  - checks out problem list with the patient 3\_\_\_\_
  - books appropriate follow-up appointment 2\_\_\_\_
  - reassures the patient appropriately 1\_\_\_\_
  - checks if there are any patient questions 0\_\_\_\_
- closes with a social amenity

## APPENDIX 1

Station B4

Rating Scale

3 marks - obtains all "must" and "should" items thoroughly

2 marks - obtains most "must" item adequately

1 mark - obtains most "must" items but incomplete

0 marks - omits several "must" items

### I. Content of the Interview

a) "must includes:

- attitudes and knowledge regarding Contraception
- gynecological history
- inquiry regarding contraindications to oral contraceptive including smoking 3\_\_\_\_\_
- past health 2\_\_\_\_\_
- medications taken at present (none) 1\_\_\_\_\_
- drug allergies 0\_\_\_\_\_

b) "should" includes:

- appropriate functional enquiry
- brief social and family history

Examiner\_\_\_\_\_

Patient\_\_\_\_\_

Total Score\_\_\_\_\_

**APPENDIX 1**

## INTERVIEWING-PATIENT EDUCATION

## CONTRACEPTION

## ACTOR'S INSTRUCTIONS

You are playing the part of Mary Jenkins, a 20-year-old secretary who presents requesting information regarding contraception. You are single and live alone in an apartment. Your boyfriend of 2 months, Bob Peterson, is not happy using condoms and, as a result, you have had unprotected intercourse on several occasions. Your last period was 2 weeks ago. Menarche occurred at age 14. Your period are regular every 28-30 days and last 4-5 days. You have never been pregnant. You have not intermenstrual bleeding, no vaginal discharge, and only mild cramps during menses for which you do not take any medications. Coitarche occurred at age 17 and Bob is your third partner. You have no history of sexually transmitted diseases. You have frequent headaches which are nuchal-occipital, bilateral, come on with fatigue, tension and hunger and are relieved by Tylenol. They are not associated with nausea or visual disturbances. You have no contraindications to the oral contraceptive pill. You believe you would be very compliant in taking the "pill" if asked. You have never had an internal examination or a Pap test. You have not been to a doctor in years.

Past History

You smoke 10 cigarettes a day and drink only socially (2-3 bee/week). You have no hospital admissions and no surgery.

Functional Enquiry

Negative aside from headaches.

Family History

Your parents, in their 50's and sister, 23, have no medical problems. You are not aware of cancer, diabetes, heart attacks or strokes in your extended family. "Everyone seems to die of old age".

Social History

You have lived in Toronto all your life and decided to get your own apartment closer to work when you got your present job at an insurance company two years ago. You get along reasonably well with your family and see them for Sunday dinner every two weeks. You have a large circle of friends, most of them you met in a commercial high school where you finished grade 12. You spend most of your spare time socializing, enjoy dancing and read romance and mystery novels.

## Acting Instructions

You are very shy and tend to answer direct questions with direct answers. Open-ended questions are rewarded with relatively extensive answers. Your knowledge of reproductive biology and contraception is quite limited. If advised to quit smoking you say you will try. You are willing to come back on another occasion for a physical examination. You will go along with whatever advice and candidate recommends but you have major reservations about every contraceptive method. You do not like taking medications and are concerned about side effects but are not aware of any specific problems associated with the "pill". You have heard that IUD is painful and condoms and diaphragms are too messy. No matter what is suggested, you ask about potential side effects. If you are asked to decide you say, "I don't know. You're the doctor."

## APPENDIX 2

### STANDARDIZED PATIENTS: THE TRAINING PROCESS

#### PREPARATION:

A number of areas require some preparation prior to beginning the actual training. These include:

The Profiles – Each profile or role must be read carefully and translated into lay terms where necessary. A fact sheet should be provided for every role, therefore, it may be necessary for the trainer to develop one. Any questions arising from the content should be noted and discussed with the physician-author prior to beginning the training sessions.

Recruiting – It is important that one attempt to match up standardized patients to their roles. This provides more reality to the OSCE. We try to select those people who most resemble the profile according to their sex, age, stature, and any specific physical characteristics as well as ethnic origin. Other qualities we look for in our standardized patients are a sense of responsibility and reliability, a professional attitude, and an ability to absorb the material accurately and quickly. A time sheet is kept for each standardized patient and is submitted to the executive assistant upon completion of the simulation for payment.

#### Training:

Two sessions of approximately 1 hour each are usually required for our OSCE focused history profiles. Only one session is necessary for the straight physical assessment profiles.

The Initial Session – It is ideal to have all of the standardized patients for a particular role attend the session together. This saves time, and more importantly, ensures standardization of the key points.

Initially, the standardized patients are oriented for the purpose of the exercise including an understanding of what an OSCE is. The purpose of this session is to familiarize the standardized patients with the profile they will be simulating. With everyone sitting together we first read the profile and discuss its content, much in the same way you would approach a script for a play. Various aspects of the profile are discussed, such as the “patient”, and all of the characteristics and details pertaining to the role.

A lot of time is spent on the actual content of the profile, making sure the people understand the terminology and the fact making sure the people understand the terminology and the fact sheet. No details are left to change. This is a good opportunity to standardize anything that could be ambiguous. Once the standardized patients feel comfortable with the profile, an interview is carried out so that they will get a feeling of how it flows. At this point questions may be asked and arrangements for follow-up meetings made. It is a good idea to space the two sessions a few days

apart but not longer than one week apart. Each person is asked to read over the role at home, to memorize the fact sheet, and to practice with it.

The Second Session – This session is held several days after the first session but no later than one week if possible. It usually lasts for one hour. The purpose is to review the material that was discussed in the initial meeting, to go over and discuss collectively any points that are unclear, and most importantly, it provides the standardized patients with an opportunity to rehearse their role in the form of a practice run.

The Trainer should be well versed in the art of history-taking. With everyone present the trainer provides an opportunity for each person to go through their profile. The standardized patient who is not playing the “patient” should act as an observer, and give feedback where indicated. It is helpful if the trainer can take the history in different ways. For example: in an organized manner to begin with, then in a slightly disorganized way. Occasionally, throwing in questions that might not be terribly relevant to the situation is helpful during training. This helps to prepare the standardized patient for the unpredictable and allows them to develop their ad libbing skills without getting too flustered. One must remind the standardized patient that the OSCE candidates will be extremely nervous and that the stress level is very high. Under that kind of pressure even the best students are likely to have some problems. The trainer should be able to prepare the standardized patients for unpredictable situations.

It is also helpful if the standardized patients have an opportunity to reverse roles with the trainer. Therefore, they take the history and observe the trainer in the “patient role”. That way, if there are any specific mannerisms or responses the trainer wants to elaborate on, these can be demonstrated.

Unless it was a very difficult role or problems arose, no further sessions were set up prior to the simulation.

### C. LOGISTICS:

The logistics involved in putting together an OSCE are complex. Because the timing is so vital to the smooth running of the exam, everything needs to be thought out in advance. These preparations can be broken down into two areas with include: a) pre – OSCE and b) OSCE day.

#### PRE – OSCE

A Letter is sent to each standardized patient at least one week prior to the OSCE. This letter contains pertinent information relating to the exam, such as exact hospital location, time and place registration, instruction regarding what to bring to the hospital (reading material was suggested) and any other specifics relevant to the day.

Each standardized patients reported one hour prior to the start of the OSCE. The orientation is then begun by the trainer. This includes any changes in the instructions previously given to them, as well as an opportunity to ask any questions they have. They are then taken on a tour of the exam area, which ahas been set up as a restricted area. Rest stations providing refreshments have been set up so patients do not have to

leave the area except for lunch which is provided for them in the cafeteria. We also show them which washrooms we want them to use. A half hour before the exam, they are asked to go to their stations where they have an opportunity to meet with the authors and the examiners. They spend approximately 20 minutes together going over their roles and discussing any pertinent issues. Ten minutes before the exam, the Group B patients are asked to leave their stations and go to the lounge that has been set up for them. The Group A patients remain in their stations to begin the exam.

Once the exam begins, the trainers circulate throughout the day between the lounge and the exam area. It is helpful to talk to the groups as they finish their stint so as to pick up on any problems. We then communicate relevant information to the opposite group. There is no opportunity for the Group A's to communicate with the Group B's except at the lunch break.

## ADDITIONAL USES FOR OSCES

*Lecia Apartaku, M.D.*

The assessment of students clinical skills is an important responsibility of the medical school faculty. The Liaison Committee on Medical Education requires institutions to develop a system of assessment which assures that students have acquired and can demonstrate on direct observation the core clinical skills and behaviors needed in subsequent medical training. In response to increasing concerns about the prevalence of knowledge-based assessments of medical student competency, leaders in medical education have emphasized the importance of methods that quantify student performance. As a result, the use of OSCEs is viewed by many as the newest and most promising technique for assessing students abilities.

Fifty six percent of 1997 medical school graduates reported their school used an OSCE (1). Besides the third year clerkship, OSCEs and standardized patients (SPs) are used throughout training. They are used or proposed for use in the first and second years, the fourth year, as part of licensure, and to teach and evaluate residents and practicing physicians.

### First and Second Year

A common use of SPs in medical schools continues to be in training students in breast and pelvic examination. Seventy five percent of schools have such a program and use it for both teaching and evaluation, although it is used primarily for teaching. The program is offered most often in the second year, often as part of the introduction to clinical medicine course. SPs are also used to teach the male genital urinary examination at 48.6% of schools, almost exclusively in the second year (2).

Fifty three percent of schools use their SPs to teach students the skills of medical history taking. The program is offered predominately in the first and second years. Feedback to the student is provided by both faculty and SPs in majority of the schools. Most of the schools give immediate feedback in both writing and oral formats. Forty seven percent of the schools use SPs for focused encounters. This category was defined as a brief history and physical examination. Focused encounters are used most frequently for teaching and evaluating second and third year students (3).

The College of Human Medicine at Michigan State University, uses an OSCE (incorporating SPs) at the end of the second year as a means of assessing students communication and physical examination skills acquired during clinical training. The year 2 OSCE has been used not as a formal evaluation gateway that students must pass before continuing on to their clinical years but rather as a formative assessment with no pass/fail consequences attached to the outcome. Third year students describe the OSCE experience as stressful but also report that it contributes to a growing sense of competence in their clinical skills. Since students consider the OSCEs to be highly instructive, they often express a desire for these exercises to take place earlier as part

of their primary instruction. Faculty say they enjoy the student contact that comes with the OSCEs, and between expected and actual student performances. Although such discrepancies occur relatively frequently, they only occasionally result in curricular changes (3).

A study of SPs in medical interviewing look at the skills of first year medical students receiving feedback primarily from faculty students. Students at the University of Oklahoma College of Medicine were video-taped to assess baseline and post-instruction interviewing. All the students attended two four-hour workshops on interviewing skills. Instruction in the groups was as similar as possible except in the matter of who provided higher for the “types of questions used” and “use of empathy” items in the SP-led feedback group (4).

#### Additional Third Year

A women’s health workshop at the University of Kentucky College of Medicine integrates new knowledge and reinforces history-taking skills necessary for women’s medical surgical care. The two-hour workshop follows an OSCE format. Small groups of two or three students with feedback regarding their interviewing and diagnostic skills at each station. Chest pain, abdominal pain, and trauma cases are utilized (5).

The standardized patient allows the student to practice with simulated emergency situations and difficult and sensitive medical conditions that you would not allow the student to work with in real patient settings. It is far better that students make their mistakes in working with a dying patient, a comatose patient, or a sexually abused patient in a simulated setting rather than in the real setting. The standardized patient, unlike the real patient, can be manipulated for educational purposes. One particularly powerful technique is the so called “time in-time out” technique. After a group of students has been working with a standardized patient for a period of time, the instructor can call “time out”. At this point the standardized patient remains in “suspended animation,” appearing in the patient role but pretending no awareness of what is happening in the room. The instructor and the students can then discuss what’s going on in their thinking, what they think might be going on in the patient, what they plan to do in the future, their interpersonal skills, and a whole variety of other things that they would not discuss in front of a real patient (6).

“The stimulated recall.” Is a powerful technique for evaluating the reasoning process of physicians when combined with a standardized patient. After the patient encounter, which is videotaped, the clinician is carefully interviewed during a replay of the videotape to probe his or her thinking during the encounter. Seeing the encounter on videotape is a very strong stimulus for recalling thoughts and ideas that occurred during the workup. I was subsequently able to expand this technique to evaluate the reasoning of medical students and also that of the residents who have difficulty with their clinical performances. This is yet another use for the standardized patient (6).

#### Fourth Year

The Clinical Practice Examination, a 17-to 18-patient examination using standardized patients that is given to all senior students when they have finished their

clinical clerkships. After a 20-minute encounter with the standardized patient, the student goes to a computer terminal to record his or her database, diagnostic ideas, and treatment plan, and to order laboratory tests. The student then has an opportunity to review the results of the laboratory tests and change diagnostic and treatment plans. This examination simulates the entire encounter with a patient from beginning to end and allows assessment of students' clinical skills in a valid, comprehensive, and powerful way (6).

The traditional role of physician examination in an OSCE is relatively passive. At Israel Institute of Technology examiners criticized the passive nature of their role. They added a structured oral examination to an OSCE. Ten 24-minute stations consisted of three parts. Part I: 12 minutes-patient encounter. II: 6 minutes-oral presentation covering findings, examination (SOE), containing 5 predetermined questions. The SOE was a reliable component of an OSCE and contributed to the overall reliability. Examiners reported a higher degree of satisfaction with the examination (7).

Four clinical progress exercises (CPEs) are given every three months during the third year of medical school, at North Western followed by a comprehensive fourth-year examination as a graduation requirement. The CPEs are formative in nature, without a grade but with the provision of immediate feedback to the student concerning his or her performance. The CPEs consist of both long (comprehensive SP evaluation) and short (15-20 minutes) stations, with SP checklists and videotapes provided to the student upon completion. Where a post-encounter challenge (e.g; the construction of a problem list for the patient) is required, a key to this is also given for self-assessment. The videotapes are accompanied by a checklist of common mistakes to guide student review, and SPs provide assistance to the students in that review. The final, fourth-year comprehensive examination is based on the preceding CPEs and requires students to demonstrate the acquisition of basic clinical skills that are requisite for graduation. Failure of this examination requires remediation (8).

### Higher Stakes Testing

The National Board of Medical Examiners' SP examination was designed to assess a physician's bedside clinical skills in a multiple station (case) exam using SPs. The assessment is targeted for use with medical students who are about to enter their first postgraduate year (PGYI) of training. Four clinical skill areas are assessed: history-taking, physical examination, communication, and interpersonal. Each case is designed to measure interpersonal skills and one or more of the remaining clinical skills. Case selection is based on a structured test blueprint that samples the content and skills that reflect current PGYI experience. History-taking, physical examination, and communication skills are evaluated using case specific, dichotomously-scored checklists containing a maximum of 25 checklist items for each case. Interpersonal skills are assessed using the Patient Perception Questionnaire (PPQ), a six-item rating scale. The case-specific checklist and PPQ are completed by the SP after each student-patient encounter. Percent-correct scores are produced for cases, skills, and the total test to reflect students' performances on the checklists and PPQs (9).

The Medical Council of Canada Qualifying Examination Part II is a 20-station OSCE consisting of ten 10-minute patient encounters and ten couplet stations, which are

five-minute encounters followed by five-minute paper-and-pencil exercises known as encounters. During the candidate-SP encounter, examinee performance is scored using a candidate's performance was satisfactory. The PEPs are scored shortly after completion of a central authority who arbitrates on problematic answers and ensures uniformity of marking (10).

The 1989-1991 ECFMG Clinical Competence Assessments study defined a process that includes 15-minutes integrated clinical encounter (ICE) with an SP. The examinee is asked to obtain a focused history, perform a relevant physical examination, and communicate the initial diagnosis and management plan to the SP. The examinee is asked to obtain a focused history, perform a relevant physical examination, and communicate the initial diagnosis and management plan to the SP. The SP then completes a checklist that documents the examinee's performance on the history and physical examination (data gathering) and rating forms that evaluate interpersonal skills and spoken English-language skills. The SP encounter is followed by a 7-minute written exercise in which the resident is asked to record on the patient's chart relevant history, physical findings, and assessment (diagnosis hypothesis) and management (immediate plans for further diagnostic workup and possible treatment). Diagnosis and management skills are assessed by clinical vignette multiple choice questions, and assessment of interpretation of diagnostic and laboratory procedures is made pictorially. The patient note also assesses written communication with the health care team (11).

### Residents

The Objective Structured Assessment of Technical Skills consists of six separate tasks that a resident performs over a 90-minute period: excision of skin lesion, insertion of a chest-tube, abdominal wall closure, control of major bleeding, stapled bowel anastomosis, and Sutured bowel anastomosis. A resident's performance at each task is assessed by a qualified surgeon using two grading methods, a task-specific checklist and a global rating scale. The task-specific checklist identifies each action that is necessary to perform the operation effectively, and the examiner indicates which of these actions are performed. The global rating scale identifies seven general rating competencies, and the examiner rates the level of each competency on a five-point Likert scale. Two versions of the OSATS have been developed, a version using live animals and a version using bench-model simulations (12).

The assessment of clinical competency is a critical component of medical education. The American Board of Emergency Medicine (ABEM) has been a leader in the development of innovative assessment techniques. In 1989 ABEM introduced a new oral recertification examination that included 3 Chart Simulated Recall (CSR) and 3 Simulation Patient Encounter (SPE) cases. CSR consists of an examiner assessing an examinee's diagnosis and management of an actual clinical case collected from the examinee's practice. Ratings are based on a structured interview conducted by the examiner and center on the chart and other supporting documentation for the case. In the ABEM oral recertification examination the examinee submits 6 consecutive cases that meet specified criteria. The 3 cases examined are randomly selected. SPE consists of an examiner role plays the patient, family members, and medical support staff while the examinee requests history, physical, and diagnostic information and manages the simulated patient (13).

Thirty residents in the first or second year in internal medicine at McGill University could compare real patients ratings for each resident in the clinic setting-a setting in which the resident physician being rated would be the primary if not sole provider of the patient's care. Second, by booking appointments for SPs in the teaching clinics, SP and Real Patient ratings of satisfaction under the same clinical conditions and during the same time period could be compared. SP ratings was a valid predictor of a resident's ranking based on RP evaluation (14).

Major deficiencies in the physical examination skills of physician-in-training have been documented. A survey of recent internal medicine graduates from Loyola University Medical Center revealed their training programs did not emphasize basic clinical skills enough to adequately prepare them for primary care practice. The musculoskeletal examination has been identified as one such area of deficiency. To address these deficiencies, an advanced physical diagnosis course was designed. As part of this one-month elective, an instructional module on the musculoskeletal examination utilizing physical therapists as patient-instructor was used. Six internal medicine residents participating in the elective received two hours of training in the examination of the knee and upper extremity. Training consisted of faculty demonstrations and ample opportunity for practice of each component of the exam. The residents received performance feedback from both faculty and two experienced physical therapists PTs, who served as patient-instructors. Related cognitive and psychomotor skills were assessed before and after training. The former was evaluated by a ten-item multiple-choice exam, the latter through the use of performance checklists. The residents performed the pre- and post-intervention musculoskeletal examinations on different PTs, who were not involved in the instruction. The PTs evaluated the residents' psychomotor performances utilizing the same checklists (15).

The internal medicine residents at Wayne State University showed a general lack of competence in delivering bad news to a simulated cancer patient. The skill items with the lowest ratings were primarily related to eliciting the patient's perspective, which may represent a weakness on the part of the residency program in teaching the residents to use "patient-centered" interviewing style. The results of this study formed the basis of a curriculum to improve residents' comfort with and skills for delivering bad news to patients (16).

There is a wide variety of the uses of OSCE'S and SP. They are used throughout the training of physicians from the first year through residency. The uses continue to expand.

**References:**

1. AAMC, The Role of Faculty Observation in Assessing Students Clinical Skills. On Contemporary Issues in Medical Education1: Number 1, October 1997.
2. Anderson, M., Stillman, P. and Wang, Y. Growing Use of Standardized Patients in Teaching and Evaluation in Medical Education. Teaching and Learning in Medicine 6: Number, 15-22, 1994.
3. Mavis, B., Henry, R. Ogle, K., Hoppe, R. The Emperor's New Clothes: The OSCE Reassessed. Academic Medicine 71: 447-448, May 1996.
4. Vanatta, J., Smith, K., Crandall, S., Fischer, P. and Williams, K. Comparison of Standardized Patients and Faculty in Teaching Medical Interviewing. Academic Medicine 71: Number 12. Pages 1360.
5. Kwolek, D., Blue, A. and Griffith, C. Integrating Women's Health Issues into Third-year Medical and Surgical Curricula. Academic Medicine, 72: Number 5, Page 422.
6. Barrows, H. An Overview of the Uses of Standardized Patients for Teaching and Evaluating Clinical Skills. Academic Medicine, 68: Number 6, Pages 443.
7. Amiel, G., Tann, M., Krausz, M., Bitterman, A. Cohen, R. Increasing Examiner Involvement in an Objective Structured Clinical Examination by Integrating a Structured Oral Examination. American Journal of Surgery 173: Number 6 Pages 546-548.
8. Winter, R. Alternative to Clinical Practice Examinations. Academic Medicine 72 Number 5 Page 444.
9. Ross, L., Clauser, B., Margolis, M., Orr, N., Klass, D. An Expert-judgement Approach to Settings Standards for a Standardized-patient Examination. Academic Medicine 71: Number 10. Supplement Page 4.
10. Reznick, R., Blackmore, D., Dauphinee, D., Rothman, A., and Smee, S. Large-scale High-stakes Testing with an OSCE: Report from the Medical Council Canada. Academic Medicine 71 Number 1 Supplement Page 19.
11. Sutnick, A. Friedman, M., Stillman, P., Norcini, J., Wilson, M. International Use of Standardized Patients, Teaching and Learning in Medicine 6: Number 1. Pages 33-35.
12. Faulkner, H., Regher, G., Martin, J., Reznick, R. Validation of an Objective Structured Assessment of Technical Skill for Surgical Residents. Academic Medicine 71: Number 12. Pages 1363-1365.
13. Solom D., Reinhart, M., Bridgham, R., Munger, B., Starnaman, S. An Assessment of an Oral Examination Format for Evaluating Clinical

Competence in Emergency Medicine. Academic Medicine 65: Number 9 Supplement Pages S43.

14. Tamblyn, R., Abrahamowicz, M., Schnarch, B., Colliver, J. Benaroya, S. and Snell, L. Can Standardized Patients Predict Real-Patient Satisfaction With the Doctor-Patient Relationship? Teaching and Learning in Medicine 6: Number 1 Pages 33-34, 1994.
15. Stirling, J., Wood, J., Lloyd, J. An Instructional Module in Musculoskeletal Examination for Residents Incorporating Physical Therapists as Patient-Instructors and Evaluators. Academic Medicine 72: Number 5 Page 453.
16. Eggly, S., Afonso, N., Rojas, G., Baker, M., Cardozo, L., Robertson, S. An Assessment of Residents' Competence in the Delivery of Bad News to Patients. Academic Medicine 72: Number 5 Page 397-399.

## THE CLERKSHIP COORDINATOR'S ROLE IN THE CREATION OF AN OSCE EXAMINATION

*Janie P. Boyer*

*Hollis W. Merrick, M.D.*

Often, the person responsible for overseeing the hundreds of details involved in the creation of an OSCE examination is the Surgery Clerkship Coordinator. An effective Coordinator should be knowledgeable about OSCE examinations, utilizing written OSCE guides and visiting other OSCE exam sites to familiarize herself with this method of evaluation. She should attend all decision making sessions, and be aware of exactly what skills and performance levels are being assessed by the OSCE, whether the exam is for evaluation only or will be utilized in the grading process.

The Coordinator will play the pivotal role in the creation of a smooth and organized examination. She usually reports directly to the Clerkship Director and acts as liaison for the entire OSCE team. The Coordinator assists with the assembly of an OSCE team, and helps to identify the duties and responsibilities of each member. The OSCE team may consist of only a few people, each performing multiple tasks or it may be a large committee with very specific duties assigned to each person. She will work closely with the OSCE Administrator, Standardized Patient Trainer and a Statistician. The Coordinator will be responsible for assisting in the recruitment and training of hall proctors, timers, data entry and other clerical personnel. She must be prepared to step into any and all of these roles at the last minute in order to guarantee the smooth running of the OSCE. The Coordinator must be goal-oriented yet flexible, with the ability to work diplomatically with many personalities while assuring that deadlines are met.

The Coordinator, while not responsible for the creation and writing of individual OSCE cases, must be thoroughly familiar with all the cases and clinical scenarios that will be presented to the student. The needs of each individual type of case, i.e., physical exam, history, x-ray, practical skill, etc. must be evaluated and organized in painstaking detail. A basic budget must be developed that will allow for examination supplies, personnel, and standardized patients for all proposed stations, as well as a myriad of miscellaneous expenses. Space must be assigned with the environmental necessities of each case in mind, and lists of the necessary personnel, equipment and supplies created.

In determining a clinical setting for the OSCE, consideration should be given to non-clinical factors such as sufficient room for registration, administration, student orientation, data management, and feedback sessions. Cases must be matched to appropriate rooms, taking into consideration the equipment needed, the type of skill being tested, whether the station is being videotaped, etc. Privacy and security are issues to be considered when choosing an OSCE site, it is best to be out of the public traffic pattern, and access to the area should be restricted to OSCE participants and personnel. It is imperative that station checklists be kept in a secure location until the

day of the examination and provided to the rater in each station in a clearly marked folder. Raters should hand checklists to the hall proctors after each student completes that station. The checklists then should be given directly to the person assisting with data entry. Students should be cautioned against conversing amongst themselves during the OSCE, and not be allowed to move on to the next station before the designated time. It is helpful to provide a break station where the students might be served refreshments.

One of the major responsibilities of the Coordinator is the computation of the OSCE grades for students. The method of grading, who does the record keeping and data analysis, and how feedback is provided to the students are issues that usually decided by the Director, but implemented by the Coordinator. Lengthy delay in production of feedback is a major source of frustration for students. Providing post-examination information as soon as possible following the exam is vital. This assures the students that the results of the examination are important and also allows them to review the information while the exam is still fresh. A wrap-up session facilitated by the exam Administrator should take place immediately following the examination, including a review of each station, allowing students to ask questions. This will enhance the learning environment and effectiveness of the OSCE.

At the end of this session, a written record of performance can be provided to each student. The methodology of this technique is provided in a separate chapter. Although it is time and labor intensive to perform the necessary tabulations for output of even basic statistical data, a quick and effective method of providing written feedback to the student is to enter data from each station into a spreadsheet program as the exam progresses. There are some excellent spreadsheet programs which allow for instant calculation and statistical analysis. This can provide valuable same-day feedback to the students. This also requires a time and labor investment on the part of the Coordinator prior to the exam, but the benefits of immediate data make it worth the effort. A computer needs to be set up in a quiet area apart from the exam traffic for a data entry station. One person should be designated to perform the data entry, with an assistant organizing the paperwork. Score sheets are taken directly from the rater at each station and taken to the data entry location. Scores from each station are entered into the spreadsheet. When all scores are entered, reports can be immediately processed and printed for each student, clearly showing that student's scores for each station as well as for the entire examination. Preparation of the input program prior to examination day is the key to successfully managing the scoring data and producing clear and informative feedback. Basic knowledge of a spreadsheet program is essential. It is necessary to insert formulas to calculate averages, totals, minimum and maximum scores. Links are created between pages, allowing a master sheet and individual station and student sheets, allowing information entered one time to be reflected in several ways on separate sheets. This is time intensive and must be done well ahead of examination day in order to provide the information needed quickly and accurately.

The Clerkship Coordinator must be detail oriented, have well-developed communication skills, be willing to devote a great deal of time and effort to the exam, and able to organize and motivate a diverse group of personnel. Most of her work is completed behind the scenes during the planning stages prior to the examination, and

following up with data analysis once the exam is finished. The key to a successful examination is preparation. The successful Coordinator will incorporate the many lessons learned as she gains experience with the OSCE, becoming a valuable and integral part of the examination process.

## EARLY FEEDBACK FOR AN OSCE

*Hollis W. Merrick, M.D.*

*George A. Nowacek, Ph.D.*

*Janie P. Boyer*

Lengthy delay in production of feedback is a major source of frustration for students and evaluators alike in an OSCE examination. It is time and labor intensive to perform the necessary tabulations for output of even basic statistical data. Of the methods of data collection and processing we analyzed, the one we chose as being most beneficial to our needs is the direct computer entry system. Utilizing this method, data can be entered into a computer concurrent with the administration of the OSCE, producing immediate primary feedback. A simple spreadsheet program such as Excel is ideal for this task.

Benefits of such a system include:

1. Students can be presented with a printout summary of their own performance in relation to their peers within minutes of completion of the OSCE.
2. Examination facilitators can have immediate knowledge of the performance of each individual student and the group as a whole.
3. Each station can be immediately assessed, identifying patterns of student performance on individual test items and/or general areas of knowledge.
4. Faculty and Resident teaching staff can evaluate whether course objectives are being met and immediately make appropriate adjustments to better address the educational needs of students.

### Method

Any system to provide student with results of their performance in an OSCE requires two separate actions: 1) collecting the performance data and entering it into an analysis system, and 2) generating a report and presenting the results to the students. There are several options for both actions which can be combined in several ways. While the system to process and report student performance in the OSCE depends on the scheme to quantify the observed performance into an evaluation metric, this paper will not discuss these schemes and will assume a checklist of performance elements has been determined for each OSCE station.

### Data Collection

There are four ways in which student performance data can be collected and entered for processing. These methods range from manual, paper and pencil recordings to direct computer entry.

The manual recording method takes written comments of the raters and adds them with the checklist of ratings from each station. The comments are from station raters are collated for each student. This method provides insight into the performance of exceptional and poor performers.

A second method of data collection and entry is the use of bubble answer sheets. This is a convenient way for raters to record their assessment of students' performance. The sheets can be overprinted with specific station checklist items; a space for written comments will also be available. This method assumes the equipment to scan the answer sheets is available, there are programs to generate the report, and the expertise to administer the system. A variation of the bubble answer sheets is the use of a flatbed scanner to read the performance data. With this technology, any recording format can be used to record numeric or alphanumeric ratings as checkboxes or handwritten comments. Forms are then placed on a flatbed scanner individually or by an optional sheet reader. The software for the system uses character recognition and graphic images to collect and store the information.

A third method is the direct computer entry of performance ratings either at the OSCE workstation or at a central computer workstation. This method of recording students' performance is becoming the preferred method because of the increasing availability of computers in the testing area and the efficiency of making the performance data available for immediate analysis and reporting. The variations on direct entry are either having one person enter all the data in a single workstation, or to have separate computers at each OSCE station and have the standardized patient or rater directly enter the data. While this method requires access to computers and expertise to develop the programs, the efficiency gained in the processing and reporting makes this the preferred method of data entry and processing of OSCE performance data.

**Table 1. Methods of data entry and resource requirements**

Entry Method	Clerical Time	Computer Expertise	Processing Time
Manual	Extensive	Minimal	Extensive
Bubbles	Minimal	Extensive	Minimal
Direct Computer	Moderate	Moderate	Minimal

Table I summarizes the methods of data entry and the resource requirements needed for each method as amount of clerical time, level of computer expertise, and the amount of time to process the information. An examination of the table shows there is not a single data entry method that requires minimal clerical time, minimal computer expertise, and minimal time.

## Reporting

There are several ways that the OSCE performance can be reported to students. These approaches range from individual feedback sessions with the clerkship director to discuss station performance, to giving students a single overall rating of performance on all stations.

At the most detailed level, students receive feedback on how they performed on each of the individual station elements. This feedback can be given either in written form as a report in a feedback meeting with the clerkship director. This approach provides the student with the most detail about their performance but requires the most time to process and feedback to the student. If the OSCE encourages raters to provide written comments about student's performance, a common strategy is to collate comments for each student from each station and include them with the detailed feedback report or meeting.

A second approach to feedback results is a report with composite scores. The selection of stations for an OSCE is often based on the objectives of the clerkship to measure specific skills, e.g., history-taking skills, technical skills, or physical exam skills. Composite scores can be computed for each skill area as the weighted or unweighted average of the common stations. The computation of skill scores should check for internal consistency of performance on the common stations.

At the most global level, student performance in the OSCE could be reported as a single score which is a weighted or unweighted sum of performance in all stations added together. This approach is the least preferred because too much station information is lost.

**Table 2.** Feedback methods, quality of feedback, and resource requirements

Feedback Method	Quality of Feedback	Clerical Time	Faculty Input	Computer Expertise	Processing Time
Personal Critique	High	Extensive	Extensive	Minimal	Extensive
Written Comments	Medium	Extensive	Minimal	Minimal	Moderate
Composite Scales	Medium	Moderate	Minimal	Moderate	Minimal
Single Overall Rating	Low	Minimal	None	Moderate	Minimal

Table II summarizes the methods of feedback, the quality of feedback obtained by each method, and the resource requirements for each method. The method of personal critique has the highest quality feedback but at the expense of high clerical time to collect and process the performance data, and considerable faculty time to discuss the results with the student. The overall rating requires minimal clerical time and can be done very quickly, but the quality of the feedback drops because all ratings are summed a information is lost. The composite rating approach provides a balance of quality feedback with minimal faculty input and processing time with moderate compute expertise and clerical time.

#### MCO Immediate Feedback System

The feedback system developed for the Medical College of Ohio surgery OSCE used a direct computer entry system with a reporting of composite performance in general skill areas. Student performance on each station is entered during the OSCE and reports are generated for distribution to students immediately at the end of the OSCE.

The computer entry system consists of recording the student performance in each station into a spreadsheet workbook. The workbook consists of three components: 1) station worksheets, 2) master data worksheet, and 3) student report worksheets. The station worksheets list the station elements as rows, and the individual students in columns. The program sums the student's performance for each of the elements and the total is paste-linked into the master data worksheet. Data pertinent to each student is concurrently paste-linked into his/her individual sheet, as is the average, maximum and minimum for each station. The individual student worksheets are single-sheet printouts, showing students how they performed on each station and the performance

of the total group taking the OSCE that day. The spreadsheet workbook is customized for the stations and students for each OSCE.

For data entry, a computer is located near the area for the OSCE and a data entry person is designated for the task. At the completion of each round of the OSCE, checklists from each of the stations are collected and delivered for data entry. There is sufficient time to enter performance data for all students in the seven-minute station cycle - five minutes stations and two minutes between stations. After the last station, students are debriefed by the clerkship director during which the data from the last cycle are entered, the data is checked, and student forms are printed. The system also prints a report of all students in all stations for the clerkship director to review.

---

**MEASUREMENT CONSIDERATIONS IN PLANNING  
AND EVALUATING AN OSCE**

*Debra A. DaRosa, Ph.D.*

*Susan Kaiser, M.D., Ph.D.*

*“Collecting (evaluation) data is much like collecting garbage.  
You must know in advance what you are going to do with the stuff before  
you collect it.”*

Mark Twain

#### Orienting Questions

1. If we implement an OSCE in the surgery clerkship, how do I ensure it is a valid and reliable examination?
2. How can you tell if the OSCE helps discriminate between weak and strong students?
3. How do we determine a score for an OSCE?
4. What kind of grade (pass/fail versus A, B, C) can be given?

There are numerous measurement issues related to any examination. It is the responsibility of the faculty to adequately address these issues to ensure the exam measures what it was designed to measure and that the results are consistent. Without appropriate attention to the various aspects of scoring and interpretation, unfair and inaccurate judgments may be made about a student and his/her performance which can have deleterious effects on the student, faculty, and the curriculum.

## MEASUREMENT PROPERTIES

To ensure it is a “good” test, the OSCE must have its measurement properties assessed. This means that evidence of validity, reliability, feasibility, and creditability should be demonstrated. The following includes brief definitions of these terms and at least one approach for determining whether or not your OSCE possesses each particular measurement characteristic.

### Validity

Validity is a quantitative expression that indicates if a test measures what it was originally intended to measure. Different types of validity can be assessed. Three options described here include predictive validity, concurrent validity, and content validity.

#### Predictive Validity

An OSCE can be said to have predictive validity if the scores are related to some actual behavior or performance of which the OSCE is supposed to be predictive. For example, an OSCE administered in the surgery clerkship designed to test “overall problem solving skills” could be validated by comparing the results of student’ overall OSCE scores to subjective performance ratings on problem solving as judged by their residency program director after one year of residency. If the scores correlate significantly, it indicates the OSCE is a valid measure of problem solving. In other words, a test has predictive validity if the score reflecting a behavior or some aspect of performance relates to a score from another measure of the same behavior at a later point in time.

#### Concurrent Validity

A test is said to demonstrate concurrent validity if there is a significant statistical association between the test results with another test or measure designed to assess the same attributes or behaviors. One approach may be to compare OSCE outcomes with results of clerkship faculty subjective ratings of the specific skills or overall skills tested by the OSCE. Agreement between OSCE test scores and faculty would be an indication of the OSCE’s concurrent validity. The main difference between concurrent and predictive validity is the time element.

#### Content Validity

This form of validity refers to the examination’s comprehensiveness or test appropriateness. The test should “test” what is expected students should know or be able to do. Content validity is determined not by a statistical procedure, but rather judgmentally. A test cannot assess every aspect of knowledge, skills, or attitudes taught during a clerkship or residency rotation. A test simply samples these areas and a good test will yield scores that are generalizable to the full set of areas taught. In other words, an OSCE that has content validity consists of a sufficient number and variety of stations that are representative of the curricular goals and objectives from which the stations were drawn and about which generalizations are to be made.

For example, let's assume we developed an OSCE designed to evaluate the operative technical skills of residents. The content validity of this exam could be established by comparing the skill areas covered in the OSCE, and the number of stations devoted to each with the skill requirements of the operating room and the relative importance of each. If the OSCE stations are representative of the latter, then the exam can be said to have content validity.

### Reliability

The reliability of an examination refers to the stability of results. It is a numerical expression of the reproducibility or consistency with which the test measures the same characteristics on different occasions. A test can be reliable and not valid; but it can not be valid without also being reliable. If you had a calculator that every time you keyed in two plus two it gave you the sum of five, it would be consistent and therefore reliable. However, it would not be accurate (valid). If the calculator reported the answer "four" each time, it would be valid and reliable. If the answer was different each time you added two plus two, it would be neither valid nor reliable. Therefore, a test that is not reliable is not a good test, even if the test demonstrates reasonable evidence of validity. The stations comprising an OSCE should be shown to be reliable.

There are different types of reliability "checks." This section will describe a few simple approaches.

### Inter-rater Reliability

If any of your stations will have different raters judging performance either through direct observation, review of essay type answers, or performance checklists, the consistency of ratings across raters should be established. This can be accomplished by having raters simultaneously, but individually, observe the same performance or score the same essay and correlate their ratings. A significant correlation of .80 or higher is desirable, but correlations of .70 have been considered acceptable. Discrepancies between faculty derived scores should be discussed to clarify scoring guidelines and provide a more precise frame of reference. A more detail account on how to enhance reliability among faculty raters is published elsewhere (1).

### Skill/Competence Reliability

The different skills or competencies your OSCE was designed to test should each be tested for reliability to ensure it is a consistent indicator of that particular skill. For example, in a study published by Petrusa (1), the following clinical skills or competencies were tested in their OSCE: interviewing style, data collection, identification of abnormalities, synthesis of data, and management plans. Each competency was tested by several stations. For example, if highly reliable, scores on stations designed to measure interviewing style would be very similar for an individual across all stations, with large variance between students. The reliability is reflected as an alpha coefficient for each competency tested. This coefficient indexes the ratio variance to total variance. A correlation of .80 or higher is desirable. The

Spearman-Brown prophecy formula can then be used to estimate the increase in stations needed to obtain an alpha coefficient of .80.

### Test Reliability

Overall test reliability of the OSCE can be established by calculating a generalizability coefficient. Keep in mind that test reliability is a function of the number of stations used. Within limits, the more stations you have, the more precise the exam. Readers are referred to their friendly university psychometricians for assistance. After reading this section on reliability, you should have sufficient understanding of reliability to communicate your needs to them more comfortably.

### Feasibility

Integrating an OSCE into a clerkship will require consideration of costs, staff, time, and logistics. Once a bank of stations has been developed and tested, the time commitment of faculty sharply declines except to refine the stations, which, like all evaluation methods, is an ongoing process.

The specific resources required to operate an OSCE depend on the length and variability of the exam. For example, an OSCE designed solely to evaluate the ability to interpret x-rays would be simpler and less extensive than an OSCE designed to assess all aspects of clinical competence. Feasibility (or resources) and logistical constraints should be considered while planning an OSCE.

### Credibility

The credibility of an exam refers to its “face validity.” It is a subjective judgment as to how “life-like” the stations are. Even the best planned and piloted OSCE will be useless if the results are not considered credible by faculty and students. Faculty should be involved as much as possible in the development and scoring of the OSCE, to enhance confidence in the results. Those faculty uninvolved in this development process should be given an opportunity to progress through the exam or, at a minimum, be invited to observe. The former is strongly recommended.

## ITEM ANALYSIS

If you want your OSCE scores to differentiate good from poor performers, indexes of difficulty and discriminability should be calculated when piloting the exam. This process is frequently referred to as an “item analysis”. The following is an approach for determining difficulty and discrimination indexes based on guidelines and examples from Tuckman (2).

After having students progress through the OSCE, sort the scores into three groups: upper one-third, middle one-third, lower one-third. Count the number of students who fall into each of three categories.

An example appears as follows:

Station	No. of Upper 1/3 Who Pass	No. of Lower 1/3 Who Pass
1	10	2
2	8	9
3	3	0
4	10	3
5	9	2
6	2	4
7	7	5
8	9	6
9	10	3
10	5	5

The discriminability index is a number that reflects the extent to which a station is responded to correctly by students who do well on the exam and incorrectly by those who don't do well. The purpose of this index is to help determine how well each station separates more knowledgeable students from less knowledgeable students. This calculation of index is:

$$\text{Discriminability Index} = \frac{\text{Number of high 1/3 who pass a station}}{\text{Total Number in both groups who pass a station}}$$

$$\text{For station 1: } \frac{10}{10+2} = \frac{10}{12} = .83$$

Proceeding with our example:

Station	No. of high 1/3 Who Pass	No. of low 1/3 Who Pass	Discriminability Index
1	10	2	.83
2	8	9	.47
3	3	0	1.0
4	10	3	.77
5	9	2	.82
6	2	4	.33
7	7	5	.58
8	9	6	.60
9	10	0	1.0
10	5	5	.50

Only five of the ten stations (1,3,4,5,9) demonstrate high levels of discriminating power in that two-thirds (.67) or more who get them right did well overall on the examination. This does not mean the other stations should be thrown out, because OSCE designers may want a balance of easy and difficult stations. OSCE designers may, however, want to ensure there are enough stations in the OSCE that do indeed discriminate among performance levels of students.

The difficulty index refers to the “easiness” of the station. Assuming 20 students took the 10 stations in our example, this index is arrived at as follows:

$$\text{Difficulty Index} = \frac{\text{Number who fail a station}}{\text{Total number of pass and fail students}}$$

$$\text{For Station 1:} \quad \frac{20-(10+0)}{20} = \frac{10}{20} = .50$$

Unlike the discriminability index, the difficult index does not differentiate among students.

Station	Difficult Index	Discriminability Index
1	.40	.83
2	.15	.47
3	.85	1.00
4	.35	.77
5	.45	.82
6	.70	.33
7	.40	.58
8	.25	.60
9	.35	1.00
10	.50	.50

From reviewing our example, it is shown that stations 2 and 8 are very easy stations because less than one-third of the twenty students failed them (below .33) and that stations 3 and 6 are difficult stations because more than two-thirds of the students failed them (greater than .67). There are obvious reasons for justifying a mix of easy and difficult stations, but the OSCE developer will need to ensure the proper balance. An OSCE with an imbalance can result in an exam that does not challenge the students nor discriminate among performance levels. Difficulty index data can assist faculty in reviewing content of stations and also has curricular implications. Our example shows stations 1,4,5, and 9 demonstrate reasonable evidence of discriminability and difficulty. Keep in mind that test reliability is affected by station difficulty levels and that stations with difficulty levels of about .50 have been found to be most useful.

Calculating these two indexes provides an overview of how many students failed each station and the extent to which the students who scored well on the exam overall did on each individual station.

Cumulative data for both indexes can help test developers select stations for an OSCE with proper balance of difficulty and discriminability.

## GRADING

There are different types of reference systems upon which grades are based. Two of these reference systems are described below. OSCE developers should discuss this decision with faculty before establishing a grading system.

The criterion-referenced grading system does not compare students to each other in their performance on an exam. A criterion or “pass” level is established, and anyone who scores at or above that level passes and anyone who scores below it fails. This is usually accomplished by using the traditional percent correct scale (0-100) where 100 is perfect performance and some value represents the minimal passing score. Pass/fail designations can be used or grades can be assigned in terms of specified ranges on the percent correct scale (e.g. 100-90 = A; 89-80 = B, etc.)

There are certain strengths and weakness associated with this system. Some experts feel that students perform better if they “compete” with each other for grades. However, if an OSCE truly samples program objectives and has demonstrated content validity, faculty may decide that those who pass 70-80% of the stations essentially have “mastered” the material and should pass.

The norm-referenced grading system is the more traditional system, in which a student’s performance is judged and reported in terms of some standard or norm derived from all the students performance on that test.

In other words, students are assigned grades relative to the performance of other students. There are several approaches:

1. Normal curve – grades are distributed in a symmetric normal fashion based on a “normal curve” distribution of scores. For example, an A=10% of class, B=20% of class, C=40% of class, D=bottom 20% of class, and F=bottom 10% of the class. We do not advocate this approach because an OSCE seldom yields a normal score distribution and there is no guarantee that the size of one class is large enough to be representative of the population of students over several years.
2. Specific class norm – this distribution assumes a certain ability level of the class. A preset distribution of grades is then established. For example, it might be set that 5% of the class fails, 30% are average, 40% are above average, and 25% are exceptional. The obvious disadvantage, like the normal curve distribution, is that 25% of the students are rated exceptional regardless of overall performance.
3. Specific class distribution – the mean and standard deviation for the class scores are determined. The grades are then based around the mean. For

example, two standard deviations above the mean is an A, one standard deviation above the mean is a B, etc.

The principle difference between criterion and norm-referenced grading systems is how the scores are used. In criterion-referenced grading, the student's performance is compared to a predetermined standard or criterion; in norm-referenced grading, the student's performance is compared to that of other students. Test developers can use a combined approach whereby a minimal pass level is established before the OSCE is administered. Those who fall below this level fail the test. For those students who meet or exceed the preset "cut-off" score, the standard deviation can then be used to determine grade boundaries. It can also be designed so that certain stations must be completed successfully in order to pass the overall exam, and advanced or more sophisticated stations can be used to distinguish quality of passing performance.

### Conclusions

Numerous measurement issues need to be explored and decided upon before administering an OSCE.

The main points of this chapter are:

1. Before being used as a performance evaluation resource, an OSCE should be tested for appropriate measurement characteristics such as validity, reliability, feasibility, and credibility.
2. There are different types of validity for which an OSCE can be tested, including predictive validity, concurrent validity, and content validity.
3. A valid OSCE station measures what it was designed to measure. A reliable station measures it consistently.
4. Item analysis should be completed for an OSCE to provide indications concerning the difficulty of each station in relation to the overall exam and how well station discriminates among various levels of performance.
5. Grading can be based on a criterion-referenced system, non-referenced system, or a combination of both. Faculty should decide in advance which system best meets their fundamental purpose for the exam.

## REFERENCES

1. Steiner, D. L. Global rating scales. In Assessing Clinical Competence. Springer, New York, 1988.
2. Petrusa E.R., Blackwell T., Parcel, S., Saydjari, C. Psychometric properties of the objective clinical exam (OCE) as an instrument for final evaluation. In Newer Developments in Assessing Clinical Competence. Heal Publications, Montreal, 1986.

Tuchman, B.W. Identifying and describing procedures for observation and measurement. In Conducting Educational Research (2<sup>nd</sup> Ed). Harcourt Brace Jovanovich, New York, 1978.