

SESSION ONE

An Overview of the Uses of Standardized Patients for Teaching and Evaluating Clinical Skills

HOWARD S. BARROWS, MD

Abstract—The author defines the term *standardized patient* (SP), the umbrella term for both a simulated patient (a well person trained to simulate a patient's illness in a standardized way) and an actual patient (who is trained to present his or her own illness in a standardized way). He first discusses the many values of simulated patients over actual patients as teaching and assessment tools in the classroom and refutes a few myths about the use of SPs. Then he recounts the origin and development of SPs over a three-decade period, beginning with his work as a neurologist at the Los Angeles County Hospital, where he trained a model from the art department to simulate a neurological patient and assist in the assessment of clinical clerks. He then describes additional roles of SPs that have developed, in-

cluding (1) their use in the Clinical Practice Examination created at Southern Illinois University School of Medicine and (2) the major use that has come into being over the last 10–15 years: facilitating the comprehensive assessment of clinical competence using multiple stations in examinations such as the objective structured clinical examination. He concludes with information about recent and current work on SPs, who are becoming more and more accepted in the assessment process, and urges skeptics not to make judgments about the value of SPs until they have experienced the technique firsthand and reviewed the literature concerning the extensive and often high-quality research about this assessment tool. *Acad. Med.* 68(1993):443–451.

A meeting such as this one concerning the use of standardized patients would have been inconceivable when I began to explore this technique 30 years ago. Originally called *programmed patients*, they later became known as *simulated patients* and now *standardized patients* (SPs). My plan is to give this overview of the uses of standardized patients in three parts. First, I will describe the value the standardized patient had for me personally when I was a full-time academic neurologist with heavy teaching responsibilities. Then I will describe its value to me later as an educational administrator working with other faculty from many disciplines over the last 15 years. The intent is to show you the wide range of usefulness this technique has in areas you may not even have thought about. Following this, I will describe the explosion of activity that has occurred in the use of the standardized patient as an assessment tool for clinical competence in multi-station examinations. I have had many fasci-

nating experiences with this tool over the last 30 years and I could go on for hours about them. I have met and worked with a number of creative and dedicated researchers and teachers who have made many contributions to the technique of the standardized patient. I can't possibly do justice to all of them here.

DEFINITIONS AND ROLES OF SPs

There are many different roles for people who are used for teaching and assessment in medical education, and not all of them are the roles associated with standardized patients. It is important to distinguish these roles and to use a consistent terminology to avoid confusion. *Role-playing*, for example, is a specific technique that has the most educational advantage for the role player. In role-playing someone is given a role to enact, such as that of a patient or a physician, and plays it on an "ad-lib" basis. Playing roles of patients can give medical students a first-hand idea what it's like to be in the dependent patient role. *Pseudo-patients* are people who have been instructed to give a complaint that would allow them to achieve ac-

cess to a health care setting or to receive care. Once they have been encountered they then behave as themselves. This has been used in the past to evaluate admission and care of patients in psychiatric settings and to evaluate physicians' abilities to educate patients about various types of medication. *Practical instructors* have been trained to teach the pelvic or genitorectal examination as it is being performed on them by the student. It's a very powerful technique for giving the students the skills and sensitivity they need to perform an accurate examination. *Patient instructors*, a contribution of Paula Stillman, are patients who have been carefully educated about their own illnesses and how they should be evaluated on history and physical. They can then give feedback to the medical student or resident about the accuracy and completeness of the workup and show how physical findings might be better elicited. This technique can extend the clinical faculty's ability to assess students' workups of patients. *Subjects* are used to demonstrate living anatomy and play no particular teaching or patient role. Unfortunately, scripted actors in medical edu-

Dr. Barrows is associate dean for educational affairs, Southern Illinois University School of Medicine, Springfield.

cational film and videotapes are sometimes referred to as simulated patients and, again, this is incorrect—they are *actors*.

Technically, there is a difference between simulated and standardized patients. Geoff Norman coined the term *standardized patient* to replace the term *simulated patient* because it underlines the major advantage of this technique: to provide a *standardized* patient problem that will not vary from student to student. I use the term *simulated patient* for a normal person who has been carefully coached to accurately *portray* a specific patient when given the history and physical examination. I reserve the term *standardized patient* as a broader umbrella for both simulated patients and actual patients who have been carefully coached to present *their own* illnesses in a standardized, unvarying way. The advantage to this is that the term *standardized patient* will not reveal to the students whether they are actually dealing with a simulated patient or a real patient in their teaching and assessment contacts.

The range of differences possible in standardized patients is also important to note. For example, in our use of standardized patients, they are carefully trained to portray an actual patient so that we can assure students that such a patient actually exists, or existed, and we can use X-rays, laboratory results, and the patient's actual course in extended teaching from that simulation. However, some faculty create totally fictional cases for their standardized patients. Sometimes standardized patients are used only for history and the findings on physical examination are given to the student. At Southern Illinois University, the standardized patient is someone who has been so carefully coached to portray an actual patient on history and physical that the simulation cannot be detected by a skilled clinician.

VALUE OF SPs

It's important to realize that standardized patients who are simulating

an illness have advantages over real patients when used in medical education settings. Some people have the misguided opinion that standardized patients are useful only when there are not real patients available to learn from in a medical education setting. Unlike real patients, the standardized patient can be available at any time and available in any setting. You can use standardized patients in classrooms and in many nonclinical areas. You can have just the patient problem you want for the teaching or evaluation of students at any particular time of the day in any location. Unlike a real patient, the standardized patient presents the same problem for all students.

The use of standardized patients avoids mistreatment of real patients when they are used for educational purposes. The standardized patient is paid to be examined over and over again by numerous students. The standardized patient is prepared for students to perform inadequately and is prepared to be used as a teaching and assessment tool. You have no concern about the student's making inappropriate remarks in the teaching situation or using poor examination techniques.

The standardized patient provides a transition to the real patient for medical students. Medical students can work with standardized patients without embarrassment about their novice status as they are learning to take histories and do physical examinations. Working with the standardized patient, they can perfect their history and physical examination techniques until they become confident. Then they are able to learn from actual patients without any distraction or hang-ups from concerns about their ability or technique. In addition, patients then feel as though they are receiving a professional service from the students and not being experimented upon by neophyte physicians.

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 things that they would not discuss in front of a real patient. The advantage of this is that the instructor can shape the students' thinking in real time instead of asking them to recall what they were thinking in their interaction with the patient after the encounter is finished. When the instructor then says "time in," the group can continue with the patient as if nothing had happened in between.

With the standardized patient, the encounter can be started over if any student feels he or she might do a better job the second time. Time and pressure constraints can be added as students grow in their abilities to handle the patient workup. Complications in the patient's problem can be added or deleted as necessary at a particular point in the student's education.

The passage of time can also be ignored with the standardized patient; thus students can encounter at one session the patient's problem at various periods of time following the first encounter. For example, they can see a patient in the emergency room and decide on what their treatment program is going to be and then see what the patient is like two hours later or a week later or several months later

and continue working with the problem. The standardized patient can be trained to quickly undergo the changes that the patient normally would undergo over time, thus allowing the students to learn continuity of practice at one setting.

As you can see from all of this, there is no intent that standardized patients should replace real patients, the intent is for them to enhance the value of real patients in the student's learning.

One of the assumed disadvantages of standardized patients is that a long training time is required to produce a high-quality simulation. This is not true; a good simulation can be produced, with someone who has never been a standardized patient before, in approximately two to three hours. There are tricks that can be used to allow the person being trained to actually take on the patient's role as a real-life role and become a very believable patient. We encourage the standardized patients not to learn the patient's role as an intellectual script but to understand what it is like to be the actual patient with the feelings and problems the patient has. An experienced standardized patient can be given a new role in an hour or so. The reports of long training times for standardized patients refer to the times required to prepare them to fill out long checklists reporting on the students' performances during the encounters.

Another assumed disadvantage is the limited range of findings that can be simulated on physical examination. The range of findings that can be simulated on physical examination is probably greater than most faculty realize. The list in the box entitled "Physical Findings That Can Be Simulated" shows the many physical findings that can be successfully simulated. This list can be augmented by using patients who have fixed physical findings as simulators.

DEVELOPMENT OF SPs

I developed this technique over 30 years ago to solve an assessment problem in a clinical clerkship in neu-

Physical Findings That Can Be Simulated

Abdominal tenderness	Hypomania
Acute abdomen	Incoordination
Airway obstruction	Jaundice
Anaphylactic shock	Joint restriction
Aphasia	Joint warmth and redness
Asterixis	Kernig's sign
Atheotosis	Kussmaul respirations
Beevor's sign	Lid lag
Brudzinski sign	Muscle spasms
Carotid bruit	Muscle weakness
Cheyne-Stokes respirations	Nuchal rigidity
Chorea	Parkinsonism
Chronic obstructive pulmonary disease	Perspiration
Coma/unresponsiveness	Photosensitivity
Confusion	Pneumothorax
Costovertebral-angle tenderness	Ptosis of the lid
Decerebrate fit	Rebound tenderness (abdomen)
Dilated pupil	Renal artery stenosis
Doll's-eye response	Retardation
Dysarthria	Rigidity
Extensor plantar response	Seizures
("Babinski")	Sensory losses
Facial paralysis	Shortness of breath
Gait abnormalities	Spasticity
Ataxia	"Stiff-man" syndrome
Hemiparesis	Tachycardia (with some SPs)
Waddling	Tenderness/rigidity on palpation
Degenerative hip	Thyroid bruit
Hearing loss	Tremor
Hematemesis	Visual loss (central, peripheral)
Hyperactive tendon reflexes	Vomiting
Hyper/hypotension (rigged cuff)	Wheezing

rology. During my last year as a neurology resident at the New York Neurological Institute, Columbia-Presbyterian Medical Center, I had two experiences that led me to believe that better assessment of clinical performance is needed and that such a technique might be useful and possible.

During that year I was assigned as a neurology consultant to the Goldbeater Memorial Hospital on Welfare Island in New York City. There I ran into David Seegal, who headed the Columbia-Presbyterian Research Unit. He had clinical clerks assigned to his medical service, and I was intrigued to learn that he would sit down and watch every single one of

the clinical clerks do a complete history and physical. This would take him approximately two to three hours with each student. Dr. Seegal said that this was important, as no one to that point had ever bothered to observe these young men and women carry out their professional task. He was able to find numerous errors in history-taking skills, physical examination skills, and thinking skills, and he provided the students with valuable feedback. The students were enthusiastic about this input and, according to Dr. Seegal, often continued to do this themselves when they were in teaching situations after graduation. I was impressed with the value of this type of assessment.

The second event that occurred in that year related to the board examination in psychiatry and neurology. As the chief resident, I had the responsibility of bringing in neurological patients from surrounding hospitals (particularly the chronic neurological patients from Montefiore Hospital) for that examination. Following the examination, the director of the Montefiore neurology service made rounds on his patients to see how they had tolerated the numerous examinations they had had to undergo during the examination. He interviewed a patient known to everybody as Sam, who had syringomyelia. When asked about the examination, Sam remarked that there had been no particular problem except with the physician who had examined him last. Sam indicated that that physician had been quite hostile and had performed a very uncomfortable neurological examination. The director said that he was sorry to hear that, but Sam said, "Don't worry, I fixed him—I put my Babinski on the other foot and changed my sensory findings." He had simulated neurological findings.

My first academic post was at the Los Angeles County Hospital and University of Southern California School of Medicine. As the first and only full-time neurologist there, I was responsible for a 100-bed neurology service, often with additional beds in the corridors. I was also responsible for the clinical clerks rotating through neurology. Clerkship rotation after clerkship rotation, the visiting neurology faculty would always report that these were all men and women with no particular problems in their performances as neurology clerks. Based on my experience with Dr. Seegal, I felt that much was being missed and that we needed a more detailed assessment of the students' performances; yet I was not able to spend the time that Dr. Seegal could with the students in his program.

At that time I was using a woman model from the art department as a subject for an audiovisual resource on the neurological examination. The idea occurred to me that perhaps we could teach this person, who had al-

ready learned a lot about the neurological examination, to be a neurological patient, and use her to assess the clinical clerks. She was coached to have a paraplegia, bilateral Babinskis, dissociated sensory loss, and a blind eye. She learned to present with the anxiety and concern of the real patient she was modeled after. She reported on the performances of the students following all encounters. This worked extremely well, and many more people were coached to be different patients for subsequent rotations. I learned a lot about the performance of students that I would never have learned otherwise. I learned about students' interpersonal skills, clinical skills, and thinking skills. The feedback I was able to provide students was powerful and was received very positively by them. In fact, the students requested a simulated patient in the middle of the clerkship so they could discover problems in their neurological workups that they might work on before the final assessment with a simulated patient.

It is accurate to say that the remaining faculty at the University of Southern California (USC) were quite skeptical of this technique and considered it of little merit. However, I received strong support from Steve Abrahamson, who helped me to develop this teaching technique and to assess its value. He has continued to be a strong advocate for the development and use of standardized patients in teaching and assessment.

This experience certainly underlines the fact that the standardized patient is not a technique to be used only when there is a lack of real patients. At that time, the neurology service at the Los Angeles County Hospital probably had the largest neurological patient population of almost any hospital in the United States. I continued to use the standardized patient in the neurology clerkship for eight years at USC.

Toward the end of that time, I was asked to train a simulated patient for Art Elstein and Lee Shulman, who were studying the clinical reasoning process of physicians and felt that this would be a useful tool in that

endeavor. During my visit with them at Michigan State, they put me through the assessment technique called "the stimulated recall." I realized that this was 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 the 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.

I spent the next ten years at McMaster University in a very stimulating educational climate. The use of standardized patients to assess clinical clerks continued there as it had at USC. This is where the use of standardized patients for clinical teaching was first augmented by the "time in-time out" technique. I used the standardized patient during the clinical clerkship to provide the students with neurological problems that were important for them to encounter, yet might not occur spontaneously on the clinical service. The standardized patient could compensate for one of the major problems of clerkship training: the faculty have no control over the patient experiences that students have. By using standardized patients, some of the common and important problems in neurology would always be experienced by the students during their clerkships.

I also found the standardized patient a valuable tool to prepare students for work in an ambulatory neurology clinic. I could use the standardized patient with the students to make certain that their neurological examination techniques were intact and that they were able to deal with a neurological problem in an efficient and focused manner.

Another valuable use of the standardized patient was in continuing

medical education. I was able to take one standardized patient into northern Ontario and provide neurology clinics for general practitioners. That one person was able to simulate at least eight to ten neurological problems, providing a whole day's clinic in important neurological conditions. This technique was extended to the American Medical Association's Bed-side Clinics in Neurology for three or four years. Here, a number of neurologists from around the country came to tutor five or six clinicians around a series of standardized patients, determining their educational needs in handling common neurological problems, and giving them feedback as they worked. These clinics were very well received by practicing physicians, who felt they were highly relevant to their needs in practice.

At McMaster we also employed the standardized patient as a research tool to examine the clinical reasoning process of the physician, building on the work of Shulman and Elstein. A five-year study with Vic Neufeld, Geoff Norman, and John Feightner analyzed the clinical reasoning processes of 31 physicians randomly selected from the community, using the stimulated recall and standardized patients. The results of that study were basic to the development of problems for use in problem-based learning, and to the development of the tutorial methods used in problem-based learning to encourage an effective and efficient clinical reasoning process. The Nervous System Unit, variously ten to 12 weeks in the McMaster curriculum, was my responsibility, and problems were designed for that unit that would allow the students to practice the clinical reasoning process of the physician, as determined by the research I just described. One of the more important problem formats in this unit was the standardized patient, which complemented the other problem formats in allowing the students to practice clinical skills and interpersonal skills as they worked through the problem in the small tutorial group.

We also discovered at this time that the standardized patient was a valuable tool for evaluating the skills and

quality of a physician's performance in his or her own office setting. The standardized patient could appear in physicians' offices unannounced and not be detected as a simulation. Physicians' performances could be compared by an analysis of their records following that patient visit and by reports from the standardized patient about the events that had transpired during the encounter.

ADDITIONAL USES OF SPs

For the last 12 years I have been at Southern Illinois University School of Medicine (SIU). There, my experience with the standardized patient changed from that of a personal tool for a neurologist with teaching and assessment responsibilities to a tool for the development of medical education programs in the curriculum. At SIU I have the privilege of working with a talented team of educators and faculty who have carried out a number of important and innovative programs. Under Michelle Marcy SIU has a coordinated standardized-patient program that serves as a resource for all four years of medical school, most clinical departments, and some residency programs. We designed and built what is called the Professional Development Laboratory, with six well-equipped examining rooms surrounding a central monitoring room in which students and residents can be observed as they work with standardized patient problems in a natural clinical setting. These rooms are equipped with video cameras and one-way windows.

Additional uses for the standardized patient have been developed here. One of the most important ones is 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 us to assess students' clinical skills in a valid, comprehensive, and powerful way. It takes each student over two days to go through this examination, and it takes us several weeks to put the entire class of 72 students through it. We have now had eight years of experience with this examination, and a report of our six-year experience is available to anyone interested.

We have extended the use of the stimulated recall, associated with the standardized patient, as a tool for evaluating medical students and residents who evidence problems in their clinical performances. The value of this tool is in elucidating whether there are reasoning problems or knowledge problems and in identifying ways in which an educational program may be devised to help the student improve. The use of the stimulated recall in this manner is very much like taking a history and doing a physical on a patient. It allows the clinical teacher to carry out an analysis of the student's or resident's performance and to determine exactly what kind of pathology, if you will, is present in his or her ability to reason through a patient problem.

SPs IN MULTIPLE-STATION ASSESSMENTS

Let me now review the major development in the use of the standardized patient over the last ten or 15 years: its use as a tool to facilitate the comprehensive assessment of clinical competence using multiple stations. In this activity standardized patients often have reporting burdens beyond just carrying out the simulations. For this kind of role, they have to have additional training to be accurate in their reports of what has transpired during the encounters with the examining students and/or residents. Here, as in teaching and research, the value of the standardized patients may exceed that of real patients, in that they can be used to assess a large number of examinees in their handling of a range of important patient

challenges.

The actual assessment of clinical competence using patients is rarely done in medical education. This was a stimulus for Seegal in 1962, to do the work that I described earlier. Paula Stillman has shown in her interviews with residents that the assessment of clinical performance is rarely carried out in medical school. If a student is watched doing a complete workup of a patient, it is rarely done more than once or twice during the student's career in medical school.

Real patients offer difficulty in assessment because they can change over time and between students. Different patients present different challenges to different students, and the faculty assessor using an actual patient may not really know the actual findings on the patient because it may not be his or her patient. Also, it is always difficult to find the appropriate patient problem for student assessment at any particular time.

All these problems with the real patient prevent faculty from establishing performance criteria ahead of time. Faculty assessments of students using real patients are usually based on individual tacit criteria without any agreed-upon standard. The students' oral and written case presentations are not a good source for the assessment of clinical competence. The faculty do not have any idea where the student may have gotten information about the patient—from another student, from an attending physician, from the patient records, from a nurse? Interpersonal skills are very rarely assessed in students' performances on the clinical services.

Unfortunately, much of the assessment of medical students' clinical performances is based on their performances on oral and written examinations. The difficulties associated with the use of real patients for the assessment of clinical performance are probably the reason for the extensive use of multiple-choice questions to evaluate clinical clerks. This is ironic, as the multiple-choice question does not evaluate what is important in clinical performance and only assures the faculty that the students are able to recall information in asso-

ciation with the presentation of a verbal stem written on a piece of paper. The multiple-choice questions are easy to give, easy to score, and produce a nice neat number that may seem scientific to faculty. The significance of the standardized-patient technique in assessment is that it can produce a valid clinical test item to assess performance that has many of the same advantages of the multiple-choice question. It is a standardized item, can be given in multiples, and can be scored in reliable and valid ways.

The concept of "authentic" assessment is becoming popular in education. This refers to assessment that directly examines the behaviors that are important. This is the advantage of the standardized-patient technique. It directly assesses those behaviors that are required in a competent clinical performance. Multiple-choice questions and oral quizzes only indirectly assess those behaviors. The standardized patient represents, then, a considerable breakthrough in performance-based assessment, providing a valid and reliable assessment tool.

One of the concerns you frequently hear expressed about the use of standardized patients for clinical assessment is "case specificity." This refers to the fact that a student's performance in dealing with one patient may be quite different from the student's performance with a second patient. You cannot predict how well a student will do on the second case from his or her performance on the first. This does not seem surprising, and I am always amazed that it is used as a concern about the use of the standardized patient. I doubt if the student's performance on one multiple-choice question will be in any way predictive of performance on a second multiple-choice question. The problem is handled in multiple-choice questions by giving a large number of multiple-choice questions so that reliability can be established. And this is exactly the case with the standardized patient; the use of multiple standardized patients in assessment gets around the problem of case specificity.

The only change in faculty attitude that is necessary is to get rid of the "do-it-in-a-day" mentality. As I have described, in the examination we give the senior students, it takes two days for them to go through 15 cases, and it takes us several weeks to put the entire class through the examination. The major standardized patient development over the last few years has been in these large-scale assessments of large numbers of students using a number of patient problems. The separate yet intertwining activities in this development represent a fascinating movement in medical education, and I will try to maintain some sort of sequence in describing them. I'm leaving concerns about psychometrics, validity, and reliability to the subsequent presentation by Jerry Colliver and Reed Williams.

Regarding multiple-station examinations, here again we have a problem in terminology. Often the term *OSCE* is used incorrectly in describing this type of examination. This acronym refers to the *objective structured clinical examination*, which was pioneered by Harden in the late 1970s and refined and disseminated by Ian Hart. It is actually similar to the practical examination that's been used in colleges and university laboratory courses produced at the medical school level. It consists of numerous stations in which the student is asked to carry out a single task or set of tasks. In one station the student may be asked to interview a patient about a headache. In another station he or she may be asked to read an X-ray. In another station, the student may be asked to examine an abdomen. There are usually stations in between in which the student is asked to answer questions relevant to the previous stations. Often, faculty members monitor the students' performances on some of these stations. The stations are characteristically 5 minutes or so long, and an entire class of students can be assessed within one day. This examination is widely used, as it represents the first opportunity to directly and reliably assess clinical performances on a large scale in medical education.

Unfortunately, almost all multiple-

station examinations tend to be referred to as OSCEs. The OSCE focuses on only one single skill or a small set of skills. The type of examination I referred to earlier, that we carry out at SIU, gives the student 20 minutes in the station encounter, and the student is asked to carry out whatever is appropriate in the workup of the patient, as would be true in the actual clinical encounter. This means that the student has to ask the appropriate questions, do what is appropriate during the physical examination, and provide appropriate patient education during his or her time with the patient. This format focuses on the student's ability to use all clinical skills and to orchestrate them in an appropriate way with appropriate priorities depending upon the problem that was presented.

The OSCE can determine whether a student is capable of carrying out a particular skill, but does not determine whether the student will use that skill with an appropriate problem. To be consistent with the OSCE terminology, this longer station that I just described could be called an OUCE, which would be an *objective unstructured clinical examination*. I would prefer calling this kind of examination a CPX, or *clinical practice examination*, as it looks at the student's ability to carry out a complete episode of clinical performance for the patient problem.

Another variety of multiple-station examination with standardized patients should be called the clinical skills examination. This is where a patient problem is presented by a standardized patient, and the faculty are interested in assessing whether the student is able to carry out all the questions on history and all the things on physical examination that ought to be performed as a response to that particular patient's complaint.

To me all these formats have a logical position in medical school assessment. The OSCE would seem to be a very important examination to do early in clinical training, perhaps during an introduction to clinical medicine course or in the first two years of medical school, to see if the student's skills on the clinical exami-

nation are intact and well performed. The clinical skills examination might be used at the end of an introduction to clinical medicine course to be sure that the student's entire repertoire of clinical skills abilities is intact. The clinical practice examination would be an examination that should be used, as we do at SIU in the Clinical Practice Examination, after clerkships, to see if the student is really capable of carrying out a complete patient encounter, applying the skills appropriate for the particular problem.

GROWING ACCEPTANCE OF SPs

In 1984, the Josiah Macy, Jr. Foundation supported a conference at SIU called "How to Begin Reforming the Medical School Curriculum." The participants were deans and educators from medical schools around the country. It was the conclusion of that conference that the most effective way to change medical education would be to change assessment methods, as assessments such as multiple-choice questions used by the National Board and by medical schools only promoted the memorization of facts and not reasoning and clinical performance. The group felt that there should be better methods for assessing clinical competence. When they were reassured that the standardized patient was able to do this, the deans at that particular meeting were skeptical but agreed to a repeat meeting four months later at which they could be shown how the standardized patient could be used for valid and reliable assessment of medical students.

That second conference consisted of a demonstration of multiple standardized-patient stations in which the participants were asked to go through them as students. The effect on the assembled deans was quite remarkable, as most became enthusiastic about this method of assessment and asked Steve Abrahamson to put together a committee to see if funding could be found to support the development of this type of assessment in medical schools. [That committee consisted of Steve Abrahamson (the

chair), Marilyn Heins, Geoffrey Norman, Paula Stillman, Dave Swanson, Dax Taylor, Xenia Tonesk, Reed Williams, and me.] This was probably the first example of the clinical practice examination, and it was the intent of the Abrahamson committee to see whether it could be further developed and extended. This committee was too far ahead of its time, and no financial support was found. Paula Stillman, along with David Swanson, moved ahead and carried out an extensive assessment program of resident skills in the New England area, and our group at SIU moved ahead and developed the Clinical Practice Examination.

In the development of our Clinical Practice Examination we have enjoyed the benefits of the clinically sensitive statistical and psychometric skills of Reed Williams, Christine Vu, and Jerry Colliver. This examination assesses our students' abilities to meet exit objectives that have been developed in this school to describe what we expect our students to be able to do by the time they graduate. After six years' experience with this examination, we asked a review panel* to give it a very careful critique. The review of this blue-ribbon panel reinforced the value of this examination in the assessment of medi-

*This committee consisted of M. Brownell (Brownie) Anderson, MEd, at that time, senior staff associate, Association of American Medical Colleges; Beth Dawson-Saunders, PhD, senior psychometrist, American Board of Internal Medicine; Daniel Klass, MD, senior medical evaluation officer and director, Standardized Patient Project, National Board of Medical Examiners; Stephen Abrahamson, PhD, emeritus professor of medical education, USC; Ian Hart, MD, FRCPC, University of Ottawa; John Norcini, PhD, executive vice president for evaluation and research, American Board of Internal Medicine; Richard Resnick, MD, MEd, FRCSC, associate professor, Department of Surgery, University of Toronto; Alton Sutnick, MD, vice president, Educational Commission for Foreign Medical Graduates; David Swanson, PhD, senior evaluation officer, Step 1, Division of Evaluation Programs, National Board of Medical Examiners; and Reed Williams, PhD, Department of Medical Education, head of evaluation programs and director of the Clinical Performance Center, University of Illinois, School of Medicine, Chicago.

cal students' competence, and they encouraged us to give it more weight in our assessment. Southern Illinois University School of Medicine is probably the first school that has used this kind of examination to assess medical students in a way that counts. For if a student fails this examination he or she must take a repeat examination. If the student fails the repeat examination, the Student Progress Committee has to review very carefully the student's test performance and also his or her performance throughout the entire curriculum, to decide whether the student should be able to graduate. (It's important to note that this small state medical school developed this examination without any external funds.)

During the 1960s and 1970s and even in the early 1980s, work with the standardized patient was a relatively lonely undertaking, as very few people seemed to be interested in the value of the technique for teaching and assessment. However, one other lone voice that appeared in the mid 1970s to advocate the use of the standardized patient in medical education was that of Paula Stillman. She developed and validated the Arizona Clinical Interview Rating Scale, an important tool for scoring competence in clinical examination. She trained mothers to simulate problems for student assessment in pediatrics. She developed the technique of patient instructors that could assess students and residents in their abilities to evaluate patients' problems. Along with Dave Swanson, she developed an extensive assessment of residents' skills in many of the New England schools. At the University of Massachusetts she also developed a program to assess the diagnostic performances of fourth-year students. More recently, Dr. Stillman has directed and coordinated a large-scale multi-center pilot with the Educational Commission for Foreign Medical Graduates using standardized patients to assess the abilities of foreign medical graduates to enter residency training in the United States. This valuable study should provide us with many insights about the use of stan-

dardized patients to assess residents. Dr. Stillman also developed the Special Interest Group on Simulated Patients for the Association of American Medical Colleges.

In the middle 1980s, the University of Manitoba joined forces with SIU and carried out an identical examination of its medical students. The two schools began developing cases in common and attempted for several years to continue with identical assessments. This was an important experience, as it demonstrated that similar examinations could be given at two different sites, where the standardized patients were trained at the two different sites, with quite comparable results.

Stimulated by George Miller and the National Board of Medical Examiners, the Clinical Skills Assessment Alliance was created, and it eventually attracted representatives from the American Board of Medical Specialties, the American Medical Association, the Association of American Medical Colleges, the Educational Commission for Foreign Medical Graduates, and the National Board of Medical Examiners. This organization was initially developed to bring together researchers who had been working actively in the area of clinical skills assessment using standardized patients. The intent was to have them address important areas of development and research that needed to be considered and to find funding for their work. This organization has facilitated an annual meeting of researchers active in the area of clinical assessment.

Daniel Klass, working with the National Board of Medical Examiners (NBME), has an extensive project working in multiple sites with medical schools to develop a pilot examination using standardized patients in multiple stations as a possible addition to the NBME examination. Dr. Klass has developed a community of trainers and interested physicians from schools around the country.

Dr. Alton Sutnick has carefully engineered a well-designed pilot project with the Educational Commission for

Foreign Medical Graduates that has considered many psychometric questions and problems associated with the assessment of clinical competence of residents. This examination combined the multiple stations of Paula Stillman with a number of OSCE stations to provide a comprehensive assessment of first-year residents' skills. Dr. Sutnick put together a think tank of people active in this area of clinical skills assessment to design the project and, using the excellent skills of Paula Stillman, carried out the pilot study. One spin-off of this pilot study has been the discovery that standardized patients can provide reliable judgements of examinees' skills in English.

The University of Illinois now has a standardized-patient facility developed by Reed Williams that may lead to a statewide consortium of medical schools that will establish clinical performance standards.

The Medical Council of Canada has brought together a number of Canadian investigators to develop a clinical practice examination that will become part of Canada's certifying examination.

The Josiah Macy, Jr. Foundation, under the leadership of Tom Meikle, encouraged SIU to carry out five workshops on the Clinical Practice Examination for the deans and senior faculty of medical schools around the country. The participants went through five standardized patient stations to witness firsthand the power of this clinical assessment technique. Ninety-one medical schools were involved in these five workshops, and many of them have subsequently become active in developing comprehensive clinical assessment programs of their own. The Macy Foundation followed up by offering six grants to medical schools that would form a consortium to develop comprehensive clinical-skills assessment techniques around standardized patients utilizing common resources. As a result, there are now six consortia made up of some 23 medical schools that are developing a variety of ways to assess clinical competence with multiple stations using standardized patients.

