
Following the Threads of an Innovation: The History of Standardized Patients in Medical Education

Peggy Wallace

Prologue

The emblem of the medical profession, the caduceus, was given to Hermes, messenger of the gods, by Apollo, the God of Medicine, who empowered the winged staff to bring peace out of conflict. Skeptical, Hermes tested Apollo's declaration by planting the golden winged rod between two fighting serpents at which point they both entwined themselves in opposite directions onto the staff, ending up facing one another in reconciliation. That totality, the caduceus, is a symbol of integration. It is the weaving of conflicting elements together into wholeness, no longer "either-or," but rather "both-and": science and art, teaching and testing, knowledge and compassion, theory and practice, technology and the person, doctor and patient, illness and recovery. It refers back to the method of Hippocrates, the Father of Medicine, and to that of Apollo's son, Asklepios, the God of Healing. It goes back to the Oath of Hippocrates taken by all who are about to embark upon the practice of medicine. In addition to a code of ethics by which the new physician will live in the fellowship of the medical profession, the Oath honors

the relationship and responsibilities of student to teacher/teacher to student.

Underneath the history that is about to unfold are many untold stories of the teachers and the students, the famous and the unknown, whose endeavors, trivial and distinguished, hold the wisdom of the serpents on the staff of the innovation. After thirty years, the standardized patient now supports that knowledge and holds in the caduceus's paired wings, the inspiration for learning from the immediacy of the human encounter. The student and the teacher coming together to discover the wisdom and the meaning of the profession that chooses them.

Introduction

The term *standardized patient* (SP) has gone through many metamorphoses as the process itself has been refined since its inception in 1963. There have been many other names attempting to describe this phenomenon: programmed patient, patient instructor, patient educator, professional patient, surrogate patient, teaching associate, and—the more generic term—*simulated patient*. What all of these terms are referring to is a person who has been



The caduceus

carefully trained to take on the characteristics of a real patient in order to provide an opportunity for a student to learn or be evaluated on skills *firsthand*. While working with the standardized patient, the student can experience and practice clinical medicine without jeopardizing the health or welfare of real, sick patients. The value is in the *experience* of working with a patient. It takes the process of learning a step beyond the books and away from reliance on paper and pencil tests. It puts the learning of medicine in the arena of veritable clinical practice—not virtual reality, but veritable reality—as close to the truth of an authentic clinical encounter as one can get without actually being there, because there is a living, breathing, responding human being to encounter.

The expression “standardized patient” was coined by the Canadian psychometrician Geoffrey Norman, who was looking for a designation that would capture one of the technique’s strongest features, the fact that the patient challenge to each student remains the same.¹ The term was adopted and generally accepted in the 1980s, when the focus of medical education research using simulated patients turned sharply toward research in clinical performance evaluation. The standardized patient offers the student an opportunity to come face to face with the totality of the patient, with his “stories,” physical symptoms, emotional responses to his illness, attitudes toward the medical profession, stresses in coping with life, work and his family—in other words, everything a real patient brings to a clinician, overt and hidden (except the necessity of actually “making the patient better”), allowing the student to go about the *process* of unfolding all that he feels he needs to know from the

veritable interaction with the patient in order to assist that person to heal.

The Threads: The Innovators

The 1960s

Today, as we enter the new millennium, the standardized patient has become one of the most pervasive and highly touted of the new methodologies in medical education. But it was certainly not always so. The standardized patient was anything but a welcome and readily accepted educational tool, especially in the early days. Its acceptance was tentatively held at arms length, criticized as too touchy-feely, too expensive, too “Hollywood.” Perhaps the last charge was made because the first simulated patient was born, if not in Hollywood, very close by, in Los Angeles at the University of Southern California (USC).

The father of this innovation in medical education and the most convincing herald in the history of the use of standardized patients is the neurologist and medical educator Howard S. Barrows, who gave birth to the first simulated patient in 1963 when he was teaching third-year neurology clerks at USC. It was not an auspicious beginning. In fact, for the whole of the time that Barrows taught at that institution, “No one else at USC was even interested in using it. . . . Nobody was even interested in trying it.”² In those early days, Barrows was often invited to speak about neurological subjects, but frequently was requested *not* to talk about simulated patients. In fact, he was seen as doing something quite detrimental to medical education, maligning its dignity with “actors.” As soon as the Associated Press got hold of the idea, it was promoted in the popular press by such headlines as:

"Hollywood Invades USC Medical School"³ and such descriptions of simulated patients as: "Scantly clad models are making life a little more interesting for the USC medical students."⁴ This made it all the more difficult for Barrows to convince his medical colleagues that the technique he was using to teach medical students was a legitimate educational tool. Their resistance persisted even after the 1964 publication of the first article on simulated patients, "The Programmed Patient: A Technique for Appraising Student Performance in Clinical Neurology" in the *Journal of Medical Education*.⁵ Barrows's coauthor on this landmark article was Stephen Abrahamson, director of the USC Division of Research in Medical Education.

The USC dean received complaints from medical schools all over the country but just decided to ignore them. However Barrows, in an attempt to legitimize the work to which he was becoming so committed, replied individually to the dean or associate dean at every single one of the complaining medical schools, sharing copies of the *Journal of Medical Education* article.

Barrows went to such trouble and persisted in using the simulated patient in his clerkship for no other reason than the fact that "it was working." Students loved the technique, and, as he said, "I was learning things about those students I would have never found otherwise."⁶ Barrows was searching for an alternative to the traditional method of evaluating students on their clinical clerkships, an unsatisfactory procedure that persists even today.⁷ When faculty got together at the end of a clerkship, Barrows remembers the conversation going something like this:

"Let me look at that student's picture. . . . Well, I think I remember him." Typically, according to Barrows, most clerks received satisfactory or better evaluations. "And I knew it was because of the way they combed their hair or how neatly they dressed or if they said 'Yes sir' and 'No sir.'"⁸ Almost never was there a student whose clinical skills were evaluated as unsatisfactory because the faculty almost never directly observed a student with patients. In fact, until the advent of standardized patients, there was no objective clinical measure by which to evaluate students.⁹

As with most innovations, several events occurred in Barrows's experience that planted the seeds for the birth of the first simulated patient. In 1959-1960, during his last year as chief resident in neurology at the New York Neurological Institute, Columbia-Presbyterian Medical Center, Barrows worked on the service of a professor by the name of David Seegal who "would sit and watch every single medical student work up a patient completely from beginning to end. That took at least an hour for each student." When Barrows asked him why he was spending so much time doing this, Seegal responded: "Nobody in medical school ever watches these fledgling medical students. . . ." ¹⁰ Barrows realized that because of Seegal's commitment to direct observation of the students, he was finding a number of skills on which the students could improve because they had not known they were doing them incorrectly.

The other seminal event occurred around the same time when Barrows, who was responsible for finding neurological patients for the board examination in psychiatry and neurology, ran into Sam, a

patient who had been the subject for this examination a number of times:

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."¹¹

Barrows's first full-time faculty position was in the Department of Neurology at USC. It was there, in the early 1960s, that he met Stephen Abrahamson, the renowned medical educator who had just been recruited to establish one of the first departments of medical education in the country. They developed a relationship that inspired the exploration of a number of innovative educational methodologies. To illustrate, shortly after he arrived at USC, Abrahamson introduced the 8mm single-concept film cartridge, one of the latest innovations in media, at a conference on medical education in Los Angeles. Barrows immediately sensed its potential. He saw how he might be able to use it in teaching the normal neurological examination by documenting each part of the exam on a series of four-minute, single-cartridge film loops. Needing a person who was completely comfortable being the subject for such a demonstration, Barrows went to the USC Art Department and hired an artists' model, Rose McWilliams. Those

8mm cartridges were used by students to learn the neurological exam on the clerkship, and as a refresher for residents who might want to review a specific part of the exam before working with a patient.

However, Barrows still had the perennial clerkship evaluation dilemma. "I had the film loop and I had Rose. And suddenly I thought, 'I wonder, if I could teach Rose, like Sam, to have a neurological problem.'"¹²

In order to evaluate the clerks, Barrows needed a case about which he knew everything—all the signs, all the symptoms. He needed a case that could be reproduced for every single clerk in exactly the same way, and he needed someone who had the time and the knowledge to record what happened in each encounter with the patient. Seegal's detailed understanding of the clinical competence of his clerks and Sam's ability to create his own simulated findings inspired Barrows to create Patty Dugger, the first standardized patient case, which was performed by McWilliams.

Patty Dugger, a paraplegic woman with multiple sclerosis, was based on a Los Angeles County Hospital patient. It is a case that is so impressive in its simulated findings that Barrows still continues to use it in demonstrations around the country. Others have found this case so rich that it has shown up throughout the years in various learning, assessment, and research projects, and is, even today, still being experienced by some students in their clinical clerkships at a number of medical schools.

After the case was developed, the question of how to actually do the evaluation arose. "Should I peek through a drape, or what should I do? I finally decided that I would make [a] checklist that Rose would fill out afterwards."¹³

Barrows monitored Rose and the students from time to time, but it was Rose who was primarily responsible for recording what happened with each student during every encounter.

So the birth of the standardized patient came out of a need for a more rigorous method to evaluate the clinical skills of third-year medical students. The methodologies designed by Barrows, from what he called a "pretty primitive" first effort, are the source of the procedures currently being refined by the National Board of Medical Examiners (NBME) and the Educational Council for Foreign Medical Graduates (ECFMG) for their anticipated clinical competence exams, to be included as part of licensure, sometime around the turn of the millennium.

Barrows was learning all kinds of things about the students on his clerkship that he knew were absent on the other clerkships, but none of the faculty were willing to change. Abrahamson had helped him legitimize the simulated patient technique outside of USC, but Barrows was still meeting with such total resistance from neurologists that he began to think about other options. "Here I am a neurologist," he recalled, "and my interest is in teaching neurology. I had a tool. Neurologists were not interested in that tool [so] I eventually became interested in working with people in other fields. I remember deliberately making that decision when I was at McMaster."¹⁴

After having spent a sabbatical year in Canada in the late 1960s, Barrows left USC in 1971 to become one of the founding faculty at McMaster University in Hamilton, Ontario, "because it was a much more understanding climate."¹⁵ McMaster had a new medical school, the first with an entirely problem-based learning (PBL) curriculum.

The 1970s

Along with his use of simulated patients to evaluate medical students, Barrows began to see the value of simulated patients in teaching and research. At the same time, he started reaching out to other practicing physicians by designing workshops to help them improve their neurological skills. Barrows's underlying philosophy in these workshops was experiential learning, learning by doing and receiving immediate feedback.

Primary among a series of such seminal workshops during the 1970s, that relied heavily on the use of simulated patients, were the "Bed-side Clinics in Neurology," sponsored by the American Medical Association (AMA).¹⁶ The day before the workshop, Barrows would bring in five prominent neurologists from around the country (who would serve as tutors for the workshop participants) and an equal number of simulated patients. The SPs were not only trained to perform several neurological cases, but they were also trained to simulate typical continuing medical education participants, such as "the one who isn't interested, the one who's always asking incredible questions, interfering with everybody else."¹⁷ This gave the neurologist-tutors an opportunity to practice and learn how to work effectively with the simulated patients and the neurological cases as well as with the types of physicians with whom they might find themselves working the following day.



Photograph taken at the Los Angeles County/University of Southern California Medical Center of Howard S. Barrows training the first "Patty Dugger" (Rose McWilliams) to produce a Babinski response when her foot is being scratched

During the workshop, each neurologist-tutor was assigned a group of five or six physicians. The challenge was "to make his group of physicians perfect by the end of the day."¹⁸ Each group would start with one simulated patient and work through a case using the "time in-time out" technique. "Time in-time out" was first used by Barrows at McMaster to enhance small group teaching. By calling a "time out," the interview was "put on hold," allowing the students an opportunity to discuss among themselves any number of issues that had arisen in the encounter as well as to brainstorm where they might like to take the interview when they went back to "time in" with the simulated patient.

During the Bedside Clinics, it was the responsibility of the tutors to detect the problems the individuals in their groups were having and then focus their next simulated patient experiences in those areas. As Barrows pointedly stated, "If you ask most physicians what they don't know, they don't know they don't know what they don't know."¹⁹ This premise is true for students of all ages, no matter how experienced they are. If the students always knew what they didn't know, they could probably figure out how to learn it on their own.

In addition to Barrows's strong belief in the efficacy of experiential and participatory learning, there is another principle in his educational philosophy that sheds light on his approach. As much as possible, Barrows believes that the student should be given an opportunity to learn in the same manner as the student is going to practice. The germ of that principle can be traced to some work that Barrows did at Michigan State University (MSU), in the early 1970s, with Hilliard Jason, Arthur Elstein and Lee Shulman of the Office of Medical Research

and Development (OMRAD).

After seeing a demonstration of the Patty Dugger case at the annual meeting of the Non-Group Society (now known as the Society of Directors of Research in Medical Education [SDRME]), Hilliard Jason, the director of OMRAD, was so impressed with the technique that he established a simulated patient program in the first two years of medical school at MSU. He designed four "difficult patient" cases for the students to experience: a hostile patient, a seductive patient, a patient from another culture, and a patient who hated physicians. During the student interviews, two cameras simultaneously recorded individual shots of the student and the simulated patient. These close-up images were electronically placed side-by-side in a single, split-screen image so that when Jason later reviewed the videotaped encounter with the student, the actions and the reactions in both the student and the simulated patient could be observed simultaneously. It was one of the first of many educational applications inspired by Barrows's simulated patient work.

Shortly thereafter, Shulman invited Barrows to develop a couple of simulated patient cases that he and Elstein wanted to use, on a new research project conceived by Hilliard Jason, to try to determine how physicians solve problems.²⁰ They were using *stimulated recall*, a technique developed by a colleague, the noted psychologist Norman Kagan. Elstein and Shulman encouraged Barrows personally to go through their research protocol using this technique. The physician was encouraged to do his usual workup of a patient while being videotaped. The researchers immediately reviewed the encounter with the physician, stopping and starting the tape, asking the physician to recall what was

going on in his mind at various moments in the encounter.

It was such an enlightening experience for Barrows that it inspired him to use stimulated recall to explore the thinking process of other practicing neurologists, and then to do the same with his residents and students. What struck Barrows when he experienced the stimulated recall at MSU was that he was not teaching in the same way that he was practicing. As he recalled:

So many faculty teach students to do a complete history and complete physical. There is no such thing. Ask every question, do everything on physical, there is no such thing. And when they get into real life, they're lucky if they have twenty minutes with a patient. And if they're in an emergency, they're lucky if they've got five minutes. You can't ask every question. So they have to know the right questions to ask.²¹

The discoveries about clinical reasoning were so potent for Barrows that he changed his approach to education, no longer teaching students the "complete" history and physical exam, the way he was taught. Instead, he provided the students with the infinite possibilities a patient provides by letting the students ask anything they wanted, either in direct interaction with a simulated patient or by building that kind of flexibility into written patient problems. In this way, the student learned what questions did—and what questions did not—have a "payoff" in relation to their hypotheses. The goal was not to memorize an exhaustive list of questions and physical exam maneuvers. It was to guide the students into learning what were the *appropriate* questions and maneuvers while helping the students think through their assumptions of what might be wrong with the patient.

While at McMaster, Barrows provided another significant contribution by expanding the potential use for the simulated patient, at the same time affirming the authenticity of the SP performance methodology. One complaint often heard about the formalized assessment of clinical skills revolves around the question of physician performance—not in the examination setting, but in the day-to-day clinical practice setting. In other words, how realistic would an SP appear to the physician unaware that the patient is simulating a case? In one McMaster study, the simulated patients were scheduled in the physician's office unannounced. The skills and quality of the physician's performance were then determined by comparing the reports made by the standardized patients, who were undetected as simulations, to the physician's office records.²² This study has inspired a number of similar studies throughout the intervening years.

Besides Barrows, there are a number of other threads that weave together to shape the history of standardized patients. Also significantly responsible for establishing the standardized patient as both a credible teaching methodology and a reliable evaluation tool is the pediatrician Paula Stillman. There have been a number of organizations as well: the Josiah Macy, Jr. Foundation, the Liaison Committee on Medical Education (LCME) of the American Medical Association, the National Board of Medical Examiners (NBME) and the Educational Council for Foreign Medical Graduates (ECFMG). These organizations, which have been primarily responsible for the firm establishment of this technique in medical schools throughout the United States, as well as Stillman and a number of other less well-known clinician educators, have all

been influenced by Barrows's continuing enthusiasm, persistence and effectiveness in advocating the use of the standardized patient over the past three decades.

Contributions of Paula Stillman

In the early 1970s, when she was the pediatric clerkship director at the University of Arizona in Tucson, Paula Stillman started using simulated mothers as a technique for teaching interviewing skills. She was inspired by work being done at MSU by another pediatrician, Ray Helfer, who had trained "programmed mothers" to give histories of common pediatric complaints. Helfer, no doubt influenced by the stimulated recall research at MSU, employed graduate students to review the tapes of each medical student encounter, then code their behaviors into some twenty-five categories.²³ Stillman found the process complicated and cumbersome. When she returned to Arizona, she was determined to develop a better instrument for teaching and assessing both the content and the process of medical interviewing—an instrument that would be based on behaviors, not abstract ideas, so that it could also be used for giving feedback to the students. The Arizona Clinical Interview Rating Scale (ACIR)—or "Arizona Scale" as it became known—was the first behaviorally-anchored Likert scale to assess medical interviewing skills.²⁴

The histories Stillman taught her simulated mothers to give were compilations of the stories of several children, often including their own, laid out in the format of a checklist. She also taught the mothers how to use the checklist, to record whether or not a given item was asked, and then to give feedback to the students on their interviewing skills.

I wasn't doing anything fancy with simulation. It was strictly common pediatric problems and, by that time, well accepted interviewing skills. The mothers would play the role of the patient and then, at the end, they'd go over the content checklist and the process checklist. In the beginning, I used to videotape everything. But [the mothers] got so good at remembering specifics when they gave feedback that I stopped videotaping.²⁵

Then, in the mid-1970s, Stillman was asked to run the physical diagnosis course. She was advised by her colleagues not to accept the position because she was too young to take what was considered to be "a dead-end job." But Stillman was drawn to education, and she saw this as an opportunity to expand the work she had started in the pediatric clerkship. She was curious to see if she could develop something for the physical exam that was similar to the interviewing scale. Stillman felt the problem with existing physical exam checklists was that they were too vague and mostly in outline form:

Examine the Heart, Examine the Eyes, Examine the Abdomen. They weren't teaching tools. By reading the checklists, you couldn't tell what behavior was expected. So I developed a physical exam checklist, with family practitioners and internists, [that] had over 200 items on it. It broke down each component of the physical exam, so when it said "Examine the Eyes," there were twenty things you had to do when you examined the eye.²⁶

Stillman found a healthy man and a healthy woman, the first "patient instructors," whom she taught to use this checklist.²⁷ Not only did they know what it felt like when each maneuver was done correctly, but they knew how to teach the student to do it properly. As she explained, "[I]f you weren't reaching up high enough

in the axilla when you were palpating the axillary lymph nodes, they could teach you how to do that. They knew nothing about medicine. They were strictly process people."²⁸ Stillman's patient instructors were not simulating a real patient, they were using their own normal bodies to teach the medical students how to do a complete, accurate physical examination using a detailed checklist designed by clinicians.

The only other physician doing anything of similar import, at that time, was the obstetrician/gynecologist Robert Kretzschmar, at the University of Iowa. In 1968, inspired by Barrows's early work with simulated patients, he developed the first "gynecology teaching associates" (GTAs). The GTAs, using their own bodies, were trained to teach students and give them feedback on how to do a proper pelvic exam.²⁹ In the beginning, the identity of the GTAs, then known as "professional patients," was obscured by covering the women's faces.

The patient's responsibility was to note the various motions and sensations of the physician's examination and compare each student's performance against these criteria. She therefore gave minimal feedback to the student on his technique. . . . The simulated patient concept, in this rudimentary form, succeeded in providing a conducive environment for instruction with a relaxed, live model, but it did little to enhance communication between student and patient or reliably evaluate a student's technical performance.³⁰

By 1972, Kretzschmar had greatly expanded the role of the GTA. No longer were the patients' identities masked. The GTA had been given increased responsibilities, including teaching the unique communication skills that go along with the practical skills of a quality pelvic examination. Kretzschmar saw that integra-



Paula Stillman with her youngest patient educator, Alexandra Roberts, at the University of Arizona School of Medicine

tion of the two skills was primary. In the open climate of the 1970s, his approach to the learning of this sensitive examination was readily received by the Ob/Gyn profession, by the women who participated in the program, and by those who benefitted from it. Within a few years, a number of medical schools started their own GTA programs, many of which are still in existence today.

Stillman eventually invited Kretzschmar to the University of Arizona to speak about his work. In the meantime, with her normal physical-exam patient instructors in place in her second-year physical diagnosis course, Stillman knew she "could guarantee that before each student entered his third-year clerkship[s], he could go through a systematic physical exam."³¹ She felt confident in the process, until one day she observed a senior medical student who was examining a patient with severe bronchiectasis:

I said to him, "What do you hear?" And he said, "I don't hear anything. The lungs sound normal." And I said, "Has anybody ever checked out your findings?" And he said, "No, but I listened all over the chest and I percussed." And I realized that I never checked

that the students could differentiate normal from abnormal.³²

This awareness inspired Stillman to search for patient instructors who had actual physical findings:

Tucson [had] a wonderful population of patients with chronic diseases who were very smart and who had made enough money that they could retire early and really didn't have much to do. I found a man with terrible bronchiectasis who had been an engineer for an astronomical observatory who couldn't work anymore. I found a woman with severe aortic stenosis. I found another one with severe asthma. I found a woman with severe arthritis.³³

These four were the original patients with chronic findings. Stillman trained these patients to use her normal physical-exam process checklist along with a new content checklist that she designed to take into account the specifics of the actual findings of each patient instructor. She then taught the patients both how to examine themselves, and how to teach the students to detect the abnormality on their own bodies.³⁴ For instance, in teaching the student, the patient would place the stethoscope properly on her own body until she could hear her own abnormal finding, then she would hold the diaphragm in place while the student listened through her stethoscope. Stillman explained that the patient would then describe in detail the features the student should be listening for: "This is a systolic murmur. First, you [will] hear S1 and then you [will] hear the murmur starting after S1."³⁵

Stillman used this method in lieu of simulation "because I had this extraordinary cadre of patients. By the time I left Arizona, I must have had seventy-five patients who had chronic stable findings

[in] every organ system."³⁶ In working with these men and women and honoring them as co-educators, Stillman commanded such respect and dedication to the goals of her teaching program that, even after she left to take a position at the University of Massachusetts, many of those same patients continue years later to hold her inspiration in their present-day work with students at the University of Arizona.³⁷

The next development for Stillman was an integration of the split: "I realized that I was doing the history and I was doing the physical and I had never really put it together."³⁸ So she began working with some of the patients with chronic stable findings, simplifying their complicated histories, and putting together for them both history and physical in lengthy 45-minute encounters to use with her residents.

About this same time, in the mid-to-late 70s, Stillman began to invite people who she felt were doing "interesting work" in medical education to present their studies in Arizona. Among them were Kretzschmar, Abrahamson, and Barrows, all of whom had developed a body of sophisticated simulations. As usual, Barrows brought a standardized patient to help demonstrate his simulation techniques. The patient enacted several cases demonstrating what most people would assume to be impossible symptoms to simulate, including a pneumothorax and a comatose patient with Cheyne-Stokes breathing who, upon stimulation, throws a decerebrate fit and stops breathing. Stillman was impressed. "I had never seen this before," she said. "I thought this was the most extraordinary thing I had ever seen. I wasn't doing work with simulation because my patients had real findings."³⁹

That simulated patient with the impressive simulated findings was Robyn M.

Tamblyn, a nurse who was working with Barrows at McMaster. Tamblyn, another enduring thread in the history of standardized patients, went on to write her doctoral dissertation on the emerging standardized patient methodology.⁴⁰ She continues to work in the field of medical education and has made significant contributions to the standardized patient literature.

In 1982, when Stillman became Associate Dean for Curriculum at the University of Massachusetts, she realized she couldn't replicate the University of Arizona program exactly as it was. "I started to do work with simulation [in Massachusetts] because . . . I couldn't find that incredible pool of brilliant patients with chronic stable disease."⁴¹ Stillman has written, published, and presented on her work from the beginning, but the truly significant contributions she has made to this field were about to come, starting with the work she did in New England in the 1980s.

The names of Howard Barrows and Paula Stillman appear like interweaving threads throughout the history of standardized patients. Where Barrows started using simulation in demonstrations and for summative evaluation, Stillman began her work using patient instructors for teaching and formative assessment. Though both started with checklists, up to the early 1980s, their two approaches to education were different. Barrows's exploration of the principle "learn-medicine-as-you-will-practice-it" led him to incorporate the less tangible elements of the clinical reasoning process into his version of problem-based learning. The veritable encounters he designed for students with simulated patients integrated cognitive learning and practical experience into the "messiness" of human interaction.

Stillman's exploration was based on improving traditional educational methods. She focused on concrete behaviors, thoroughness in the basic skills of interviewing, medical history-taking and physical exam to assure that students were prepared for their required clinical rotations. However, for both Barrows and Stillman, the simulated patient became the vehicle by which they were able to investigate their clinical education insights, realize the significant accomplishments of those explorations, and, in so doing, hold the threads until the climate was conducive for others to weave in their own investigations.

The 1980s and 1990s

In 1981, Barrows left McMaster to become Associate Dean for Education at Southern Illinois University (SIU) School of Medicine. It was there that his emphasis in using the simulated patient changed from "a personal tool for a neurologist with teaching and assessment responsibilities to a tool for the development of medical education programs in the curriculum."⁴² In June 1984, on the anniversary of the tenth commencement of the SIU School of Medicine, Barrows and a number of others "felt that this milestone should be celebrated in conjunction with another event affirming the school's mission. . . . The increasing national concern about curricular abuses suggested that a conference focusing on curricular reform issues would be appropriate."⁴³

This invitational conference, "How to Begin Reforming the Medical Curriculum," co-sponsored by the Josiah Macy, Jr. Foundation and the SIU School of Medicine, ignited the adoption of standardized patients, exploding their use in medical schools across the country. Up

to that point, simulated patients were seen, by all but a few fervent advocates, as not much more than an interesting educational device. This conference provided the impetus to begin scrutinizing the efficacy of evaluating clinical competence by using standardized patients in multiple-station, performance-based assessments. The standardized patient examination was beginning to be seen not only as a valuable tool for individual student assessment, but, more potently, as the means for instigating curricular change in medical education.

In an effort to convince deans and associate deans of the usefulness of standardized patients, the Macy Foundation supported a number of follow-up experiential, standardized patient demonstrations. The first of these, "Newer Approaches to the Assessment of Clinical Performance," occurred in October 1984, when the attendees of that first invitational conference, which had taken place four months earlier, were invited back to SIU for a hands-on, multiple-station standardized patient demonstration that took place in the first fully-equipped, dedicated simulated clinic in the country. Designed by Barrows, this Professional Development Laboratory, as he called it, became the model for other schools as standardized patient programs grew and the need for dedicated clinic space became a reality.

No other demonstrations were funded by the Macy Foundation until just before Thomas H. Meikle became president. Meikle, himself a former dean of the Cornell University School of Medicine, knew that if this new standardized patient methodology were to have a chance of getting a toehold in medical schools, it would be the deans of the schools who would need to be convinced of its value. Meikle put together a blue ribbon com-

mittee chaired by David Rogers, to study clinical medical education, particularly performance-based assessment. One of the six recommendations for medical school faculties that came out of that conference was to "require medical students to pass comprehensive performance-based clinical examinations" before graduating.⁴⁴ Following this conference, his vision of the importance of clinical skills assessment reinforced, Meikle "began a process basically to educate deans and to educate my board."⁴⁵ The 1960s education mantra "Evaluation drives the curriculum" guided Meikle's efforts. He, along with many other medical educators, knew that if one wanted to make a difference in the way medicine was taught, one needed to change the way it was evaluated. Meikle has consistently held the vision of direct observation of clinical skills as a critical element in the education of medical students, even when his Board of Directors, which included a number of physicians, expressed uncertainty about the efficacy of performance-based assessment.

The next Macy support was received by Abrahamson at USC. In 1987 at Asilomar, California, he led the effort to win the support of medical school deans with another multiple-station, hands-on demonstration using standardized patient cases from SIU. This time the demonstration was solely for deans of the western regional medical schools. Following the success at Asilomar, Barrows modeled similar demonstrations for deans at medical schools in the other three regions of the country over the next several years. Out of these five Macy-funded participatory demonstrations, involving ninety-one medical schools, came enthusiastic interest in standardized patient examinations as a

potential, viable solution to the evaluation of medical students' clinical competency. It came from the majority of medical schools in the United States; and it came from the highest academic level, from medical school deans, whose written responses to these demonstrations were sent to the Macy Foundation as letters requesting financial support to explore performance-based assessment at their own institutions.

Meikle wanted to respond without delay to the deans' interest, but he was still dealing with some members on his own board who were "very unconvinced" because they felt that medical students would reject these "surrogates who are pretending." He contacted Mt. Sinai, a medical school in close physical proximity to the Macy Foundation in New York City, to see if the dean would be interested in starting a standardized patient program. Meikle's idea was to provide a firsthand experience for his board of trustees, similar to the demonstrations that had proven to be so convincing to the medical school deans. By 1990, as a result of a burgeoning standardized patient program, the Morchand Center for Clinical Assessment was built at Mt. Sinai. Meikle then persuaded the trustees of the Macy Foundation to hold one of their meetings at the new center. "That," he recalled, "was the way I convinced, particularly the physicians and educators on the Macy board, to set up the consortia."⁴⁶

In 1991 and 1992, the Macy Foundation awarded grants to support six consortia with the expressed purpose of building cooperation among schools as each consortium developed a capability for designing and utilizing a standardized patient Clinical Practice Examination (CPX) for their students. The Macy Foundation also had the foresight to establish an umbrella

Institutions Participating in Consortia, 1996
Consortium contact is italicized
 Educational impact of *Macy Project Affiliated Consortia (EMPAC)*
 coordinated by *Southern Illinois University School of Medicine*

Josiah Macy, Jr. Foundation-Sponsored Consortia (28 schools)
Gulf Coast Regional Consortium for Assessment of Performance:
University of Texas at Galveston, University of Texas at Houston, Baylor College

Metropolitan New York Center for Clinical Competence: *Mount Sinai School of Medicine, New York Medical College, Albert Einstein College of Medicine, Cornell University, New York University, State University of New York (SUNY) at Brooklyn, SUNY at Stony Brook, Columbia College of Physicians and Surgeons*

North Carolina Medical Schools Consortium: *University of North Carolina, Bowman Gray, Duke University, East Carolina University*

Northwest Consortium for Assessment of Clinical Performance:
University of Washington, University of Nevada, Oregon Health Sciences University, University of Colorado

Southern California Consortium for Assessment of Clinical Competency: *University of Southern California; University of California, Irvine; University of California, Los Angeles; University of California, San Diego; Loma Linda University*

Upstate New York Clinical Competency Center: *Albany Medical College, SUNY at Buffalo, SUNY Health Sciences Center at Syracuse, University of the State of New York Regents College Nursing Program*

Independent Consortia (13 schools)
Chicago Clinical Skills Consortium: *University of Illinois at Chicago, Finch University of Health Sciences/Chicago Medical School, Loyola University, Northwestern University, University of Chicago, Rush Medical College*

New England Consortium: *University of Massachusetts, Boston University, Harvard University, Tufts University, Brown University, Dartmouth University, University of Connecticut*

organization which was responsible for regularly bringing together the leaders from the six consortia in order to share information and to ensure that the educational impact of consortia activities would be measured and documented. That organization, EMPAC (Educational Impact of Macy Project Affiliated Consortia), was established at the SIU School of Medicine under the direction of Howard Barrows. Along with the twenty-eight medical

schools involved in the Macy Consortia,⁴⁷ two independent consortia were formed: the New England Consortium by Stillman in the 1980s⁴⁸ and, most recently, the Chicago Clinical Skills Consortium by the psychometrician and medical educator Reed G. Williams. Schools in these eight consortia, along with several other individual schools, represent almost one third of American medical schools that have been or are currently engaged in the development of a required performance-based clinical assessment of their students.⁴⁹

Integrating the Wisdom of the Staff

The Association of American Medical Colleges (AAMC) and the American Medical Association (AMA) were responsible for metaphorically digging the post-hole for the rod of the caduceus through their interest and their activities in supporting the firm establishment of standardized patient methodology into medical school curricula. First came the criticisms and recommendations for American medical education in the 1984 report *Physicians for the Twenty-First Century: Report of the Panel on the General Professional Education of the Physician and College Preparation for Medicine* (commonly known as the GPEP Report).⁵⁰ Then, the AMA Liaison Committee on Medical Education (LCME) formally incorporated into its accreditation standards the directive requiring that each medical school "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."⁵¹ Finally, two reports were published by the AAMC, which sponsored the 1992 "Consensus Conference on the Use of Standardized Patients in the

Teaching and Evaluation of Clinical Skills."⁵² As a consequence, the interest in standardized patients as a means of assessing clinical competency grew, not just in the rarified air of the ivory tower academic or the medical education theoretician and researcher, but, more importantly, at the grassroots level of individual medical schools. A 1993 AAMC survey sent to all 142 North American medical schools requested information on use of standardized patients. Of the 138 schools responding, 111 reported using SPs for both teaching and evaluation and thirty-nine of those schools were using standardized patients in a comprehensive examination to assess clinical skills before graduation.⁵³

Weaving the Fabric: Standardized Patients and Performance Assessment

The major focus of medical education research involving standardized patients from 1984 to the present has been performance assessment. It is the emphasis on performance assessment, in its many variations, that has given face validity to the use of standardized patients and encouraged the wide spread acceptance of this educational innovation. Because of the emphasis on evaluation during this period, there has been a shift away from all previously used terminology, in preference for the almost exclusive use of the expression *standardized patient*. Though one often still does hear the term *simulated patient*, it is primarily used to refer to the SP as a more generic teaching and learning tool, rather than exclusively as one for evaluation.

It is perhaps also valuable to clarify the differences between the two types of performance-based assessments: the Objective Structured Clinical Examination (OSCE) and the CPX. The OSCE was introduced in Scotland in the mid-1970s by

Ronald Harden of the University of Dundee.⁵⁴ In the intervening years, the OSCE has been refined by Harden's Scottish colleague Ian Hart, currently on the Faculty of Medicine at the University of Ottawa. Hart has been responsible for the introduction of standardized patients and the OSCE into many specialty examinations of the Royal College of Canada.

The OSCE tests specifically defined single skills. OSCE station instructions might direct the student to perform a chest exam, take a blood pressure on a real patient, take a substance abuse history from a standardized patient, start an IV on a plastic model arm, read an X-ray, or interpret lab results. In Barrows's words, the OSCE assesses the skills of the examinees by "taking a biopsy" of their clinical ability. The OSCE generally relies heavily on real patients. It may or may not incorporate standardized patients. Station length is usually short (four to ten minutes), depending on the complexity of the individual tasks comprising the exam. And in an OSCE, it is most often faculty who observe and rate the student's performance.

On the other hand, the CPX is designed to give students the opportunity to perform with a standardized patient as if they were practicing clinicians in an actual encounter. Students must rely on their own clinical judgment, responding in whatever way seems appropriate based on the patient complaint. The CPX is designed to assess the whole clinical process including history-taking, appropriate focused physical examination, patient education, and interpersonal skills. CPX stations are generally a minimum of fifteen to twenty minutes in length. The cases are portrayed by carefully trained standardized patients. And in a CPX, it is the standardized

patient who records the examinee's behavior on a checklist after each encounter. Barrows summarized the differences succinctly:

This [CPX] 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.⁵⁵

Since there has been a tendency to call all multiple-station, performance-based exams OSCEs, Meikle convinced Barrows of the importance of coming up with a name that distinguished the OSCE from the type of clinical exam that Barrows was engaged in at SIU. It was at Meikle's urging that Barrows searched for a name: "When you hear it you know what it is. It's a Clinical Practice Examination. You're examining the student in a clinical practice situation—complete interviews with a series of patients, like in a practice."⁵⁶

After Barrows's first multiple-station standardized patient demonstration at SIU in October 1984, the ten invited deans empowered another of the conference participants, Stephen Abrahamson, to convene a committee to further the development and medical school use, of this type of performance assessment, and to find funding to support that effort. This task force, known as Project Mousetrap, was ahead of its time in 1985.

Two activities were explored by Project Mousetrap: "The Snake Oil Project"⁵⁷—"traveling road shows to sell schools on the use of standardized patients in the assessment of student clinical performance"—and "Building a Better Mousetrap"—the

**Number of Articles Published on
Simulated/Standardized
Patients, 1966–1996**

MEDLINE and PSYC INFO	Citations Found
1966–1976	5
1977–1986	33
1987–1996	149

establishment of cooperation among several schools to determine graduation objectives, to design examination blueprints, and to develop quality cases, psychometrically sound checklists, and standardized patient training protocols that would assure test reliability. According to Abrahamson, "The Task Force agreed that it was really interested in the Better Mousetrap and turned its attention to the development of this project."⁵⁸ Ironically, though no funding was found before the group disbanded, a few years later both of the concerns of Project Mousetrap became the missions of a combination of other organizations: the Macy Foundation through its support of the "travelling road shows" for deans and through the establishment of CPX-based consortia; and the National Board of Medical Examiners through its ongoing valiant efforts to establish a performance-based, clinical-competency examination as part of Step II of the United States Medical Licensure Examination (USMLE).

Besides Abrahamson, Barrows, and Stillman, and representation from the AAMC, the committee consisted of a number of psychometricians: Geoffrey Norman of the Department of Clinical

Epidemiology and Biostatistics at McMaster; David B. Swanson of the American Board of Internal Medicine (ABIM); Dax Taylor, Vice President for Test Development at the National Board of Medical Examiners (NBME); and Reed Williams of the Department of Medical Education at SIU. The psychometric aspects of performance-based examinations were already becoming the principal focus of research. The major measurement concerns were the usual ones: reliability and reproducibility of test results, validity, feasibility, scoring, and reporting. However, the unique demands of this new test modality—using multiple standardized patients who perform the same case within a site or across sites; multiple standardized patient raters using the same checklist; number and length of cases needed for a reliable measure; criteria to determine "clinical competence"; and many more such concerns—challenged the creative thinking of psychometricians, most of whom had previously worked in the more cognitively-pure area of multiple-choice examinations.

By the end of 1985, it was clear that the climate among the foundations that traditionally fund medical education projects was not quite ripe for the ambitious consortial nature of Project Mousetrap. However, under Abrahamson's leadership, Project Mousetrap did inspire a number of important individual efforts from among members of the task force that, in turn, spawned research by other investigators beyond the group, finally galvanizing critical financial support from some of the very organizations who failed to see the landmark nature of what the Better Mousetrap group had attempted to accomplish. These generative research studies have contributed significantly to the

improvement of CPX psychometrics and to the dissemination of the clinical practice examination across a broad range of medical schools.

A decade after Barrows and Williams mounted the first clinical practice examination at SIU in 1985,⁵⁹ two key review articles finally gave legitimacy to large-scale standardized patient-based examinations: "Assessment of Clinical Skills with Standardized Patients: State of the Art" by Cees van der Vleuten and David B. Swanson (a review of psychometric research)⁶⁰ and "Use of Standardized Patients in Clinical Assessments: Recent Developments and Measurement Findings" by Nu Viet Vu and Barrows.⁶¹ With the publication of these reviews focusing on the psychometric aspects of large-scale clinical assessments using standardized patients, it was no longer necessary to justify the essential reliability, validity, and feasibility of these types of tests. One no longer had to apologize to skeptics of standardized patient-based examinations, one merely had to refer to these reviews.

During the same time, Paula Stillman and David Swanson teamed up, working together for about six years with funding from the American Board of Internal Medicine, which was searching for a better performance-based method of assessment for certification than its Clinical Evaluation Exercise (CEX), which involved a single faculty observation of a resident with a patient during the first postgraduate year of training. They completed the first-of-its-kind, multi-institutional study of the clinical competence of internal medicine residents using standardized patients,⁶² and eventually established a CPX-type examination for fourth-year students at the University of Massachusetts and several other medical schools in the New England area.⁶³

Simultaneous with these efforts came the National Board of Medical Examiners' Standardized Patient Project exploring a "high stakes" clinical assessment component for licensure, the Macy Foundation support of medical school consortia, and the Educational Council for Foreign Medical Graduates' (ECFMG) pilot project to assess the clinical competence of graduates of medical schools outside of North America who were in U.S. residency programs.⁶⁴ All of these efforts were aimed at finding a viable method to accurately assess clinical competence, the consequence of which was the further development, refinement, and ultimate acceptance of standardized patients as that vehicle.

One of the other threads in this narrative is Daniel J. Klass, the director of the NBME Standardized Patient Project. His coming to the National Board of Medical Examiners interweaves with a number of other threads in the history of the growing use of standardized patients. In the early 1980s, Klass was the Associate Dean for Medical Education at the University of Manitoba. He, like everyone else who is required to write "dean's letters" of reference for students embarking on postgraduate work, was dependent upon written evaluations from faculty who often disagreed about the quality of the student in question. "I started looking into the literature on clinical evaluation and found very little to go on," he said. "The literature was depressing. Whatever you saw said exactly the same thing, clinical evaluation was not very good."⁶⁵

Around the same time, a nurse by the name of Robyn Tamblyn moved to Manitoba. According to Klass: "She came into my office wondering if there was any work that she could get in the Department

of Medical Education and I didn't know who she was. I'd never heard of Robyn Tamblyn. I'd never heard of Howard Barrows. I found out pretty quickly."⁶⁶

Robyn Tamblyn, of course, had been one of the first simulated patients trained by Barrows at McMaster. Her introduction to the world of standardized patients came while she was working with him as part of a neurological patient care team. When one of Barrows's simulated patients became too pregnant to perform at a meeting of the Association of Neurological Professors, Barrows convinced Tamblyn, who had never heard about the technique before, to become an SP. According to Barrows: "She became one of the best SPs I ever worked with and in a very short while became a trainer of SPs and set up the first organized SP program at Mac. That means that when she presented herself to Klass, it was not just as an SP, but as an SP trainer and program director."⁶⁷

Tamblyn encouraged Klass to think about initiating a standardized patient program and introduced him to Barrows at the next AAMC meeting. Klass visited SIU shortly thereafter, went through Barrows's classic hands-on experience with a standardized patient, and, promptly decided to do a two-site CPX project together with Barrows. "Out of the blue, [we] created a standardized patient program that piggy-backed onto Howard's, in that we started by just saying, 'We're going to do an exam with SIU.'"⁶⁸

SIU initiated the project by supplying the cases, the faculty demonstrations, and the expertise. Within a few months, Klass and Tamblyn had mounted an examination for all the students at the University of Manitoba using standardized patients who had been trained locally. Out of the two-

school examination experience, between 1985 and 1987, came a number of interesting findings regarding portability of cases and standardization of SP performances across sites.⁶⁹

Klass's early connection with Barrows was significant in influencing the work he was about to begin at the National Board. Barrows, he said, "did with us what we have since tried to model with many other schools. The best way to learn how to do standardized patients is to do it along side of someone who has already done it before. It's [the] apprenticeship system."⁷⁰

The NBME has never worked in isolation. It has always seen the success of its licensing mission dependent upon collaboration, formal or informal, with its support coming from the State Federation of Medical Boards, the medical schools, and their faculty. This underpinning continues to guide the philosophy of the Standardized Patient Project in its work with medical schools, as well as with other organizations exploring the same territory.

From the beginning of his presidency at Macy, Meikle stayed in conversation with the NBME, whose own issues helped him define the next steps for the Macy Foundation. "Basically what I thought they [NBME] were concerned about was how they were going to test 16,000 students. And I thought this was a legitimate issue. So it seemed to me, the question was: 'What could schools do with [a] relatively minimal amount of money? And what could they do by *sharing*?' And that was the consortial concept."⁷¹

As it has evolved, the success of the Standardized Patient Project has depended, in some measure, on the very existence of the Macy consortia. Over the last five years, the twenty-eight Macy schools have provided the National Board

with a majority of its sites for piloting everything from logistics, to trainer education, to case performance. The Macy Foundation has begun to establish grassroots acceptance of standardized patient examinations at the medical school level. In fact, the dependence of faculty on the data from such examinations, which have been in place for several years, has laid the groundwork so necessary for the NBME to function. In Klass's words: "A standardized process that can be used as part of licensure across the whole country will only work if standardized patients become part of the culture of medical schools using standardized patients for their own purposes, not to meet the needs of licensure, but for their own evaluative purposes, for their own teaching purposes."⁷²

There is an ongoing struggle to integrate the focus of the medical schools—which is to educate its students to the highest standards of excellence—and the concern of the National Board to establish "minimal national standards of clinical competence" for licensure. But, in many ways, there is a recognition by both parties that each needs the other. In order to maintain the existence of the standardized patient teaching and assessment programs at the medical school level, many schools need the pressure of the NBME interest in the clinical practice examination. This is particularly important in the budgetary atmosphere of the 1990s, when what happens in medical education is often defined by "fiscal exigency." In such a climate, no matter how powerful the data, an innovation that is just becoming established can easily be cut from the medical school budget. On the other hand, the NBME needs the continuing fertile ground of grassroots, standardized patient

programs because it is the medical schools that likely will provide the professional sites where the clinical competency licensure examination will be delivered when it is ready. These mutual needs are bringing integration to traditionally competing forces. It is the wisdom of the serpents—the essence of the caduceus.

Separate but interwoven with the NBME Standardized Patient Project is a similar effort by the Educational Council for Foreign Medical Graduates (ECFMG). Here, again, a number of threads have come together. Alton I. Sutnick, former dean of the Medical College of Pennsylvania (MCP), was one of the original deans invited to the Macy/SIU invitational conference and demonstration in 1984. Abrahamson remembered Sutnick's response, more than any one else's, to his first experiences with standardized patients at that demonstration: "Each time he came out of a case, his eyes were as big as saucers because he could see the power of this thing."⁷³ Sutnick's own remembrance of that experience corroborated Abrahamson's: "Boy, was I impressed! I thought it would be something like a regular role playing, but it was so much more. I could really see what he [Barrows] was talking about so enthusiastically. He had said: 'You'll never appreciate it until you experience it yourself.' . . . It was during that day that I saw that this really did have the potential for assessment, and that it ought to be promoted."⁷⁴

Within a short time, Sutnick invited Paula Stillman to speak with his faculty at MCP, found clinic space, and appointed the team that established the first standardized patient program in Philadelphia.

A couple of years later, Robert Petersdorf, then president of the AAMC, invited Sutnick to sit with him in one of

the two AAMC seats on the ECFMG board of trustees. ECFMG, whose mission is to certify foreign medical graduates for practice in the United States, was an organization with two attractions for Sutnick, working with internationally trained physicians and exploring the assessment possibilities of standardized patients. Sutnick remembered:

The ECFMG board had already begun looking into how to test clinical skills with some studies [that had been] done in the mid-'80s.⁷⁵ These involved taking histories and doing physical examinations on real people. There were a lot of psychometric problems, so that the board wasn't ready to accept that direction. They appointed a new committee to review [that work] and decide what to do.⁷⁶

Because of his experience at MCP, Sutnick was asked to serve on that committee. Through a series of circumstances, Sutnick recalled, "I found myself in the role of planning and making recommendations on what ECFMG should do." Within a short period of time, the new president of ECFMG asked Sutnick if he might be interested in being a candidate for vice president. Years before, Sutnick had enjoyed working collaboratively with basic scientists from other countries and had been interested in the concerns of foreign medical graduates since his work in the 1960s with the Philadelphia County Medical Society, where he provided "hospitality for foreign physicians, including foreign medical graduates who were taking residencies in Philadelphia." "Little did I know," he said, "that some 25 years in the future I would become vice president of ECFMG."⁷⁷ Sutnick left MCP to take this new position in 1989.

As vice president, he became responsible for clinical skills assessment. The first

thing he did was bring together a distinguished group of educators who had been working with standardized patient-based performance assessment, including Howard Barrows, Paula Stillman, Ian Hart, and two psychometricians—Miriam Friedman, who had been consulting with ECFMG from the University of New Mexico, and John J. Norcini of the American Board of Internal Medicine.

Because of the communication through medical education literature and the cross-fertilization taking place directly among the principal players, the various approaches to standardized patient-based performance assessment had essentially converged. According to Sutnick, the ECFMG clinical assessment committee "suggested that Stillman be the active person working with us. She had developed a system, a process. She had a group of people who could contribute as collaborators. And that was crucial."⁷⁸

Stillman remembered how quickly the committee was able to get the pilot ready: "We set up four centers around the country. And I used cases that I knew. We trained patients at each of the sites. And it was done. . . . We put the exam together in a year."⁷⁹ So between 1990–1991, two pilot studies were mounted by the ECFMG using the same standardized patient cases that had been developed in New England for the fourth-year medical students' clinical examination.⁸⁰

A race to be the first to adopt a high-stakes, large-scale standardized patient clinical competency examination for licensure came to the forefront in the early 1990s. The Medical Council of Canada (MCC), under the direction of Richard K. Reznick, was exploring the use of standardized patients for a national OSCE-type certification examination. Reznick's

interest in standardized patients grew during the year he spent working with Barrows at SIU while earning a master's degree in education. In 1993, under Reznick's direction, the MCC became the first organization to implement a national standardized patient-based performance assessment as a required part of the licensure examination.⁸¹ In 1994, the ECFMG authorized the Clinical Skills Assessment as part of its certification process.⁸² And in 1995, the NBME endorsed the use of a standardized patient examination as part of USMLE Step II with implementation to be within the next four to seven years.⁸³

Along with the growing use of standardized patients in North America, there has been a corresponding interest in standardized patients internationally. In 1985, Ian Hart and Ronald Harden organized the Ottawa Conference on Assessing Clinical Competence. This biannual forum has become "the largest regularly held international conference on medical education."⁸⁴ Over the years, standardized patients have gradually been incorporated into the curricula of other medical schools around the world through the important work of such individuals as Ronald Harden in Scotland, David Newble in Australia, Cees P.M. van der Vleuten in The Netherlands, and Nu Viet Vu, who recently moved from SIU to Switzerland. Paula Stillman introduced the standardized patient to China; at the same time, the ECFMG has shown an interest in transporting the standardized patient methodology to a number of other countries with the dream of ultimately establishing global clinical standards. Towards that effort, the ECFMG, along with the World Health Organization, has assisted Israel, Spain, Russia, the Ukraine, and Brazil to mount standardized patient examinations that are

designed for each of their country's needs.⁸⁵ In little more than three decades, the use of the standardized patient in medical education has grown dramatically from its modest beginnings in California as a pedagogical novelty to the global phenomenon that it is today.

Conclusion

Standardized patient methodology is no longer in question. Yet, in our rush to quantify and establish its efficacy, a new question emerges. Have we not forgotten how much potential there is for the standardized patient in other areas—those wings on the caduceus? Assuredly, there is much left to be done in psychometrics, especially in the area of validity. Are we, in truth, assessing what we think we are—clinical competence? Is the design of the checklist, are all those details, really what we care about, or is there some other way to look at this? Now as much as ever, our creative instincts are called for.

As managed care is in ascendance, so might there be even more creative ways to use the standardized patient. The "patient instructor" might become a necessity rather than a luxury—and standardized patients might be even more extensively needed for clinical learning and self-assessment as the pool of teaching faculty dwindles. And what about the practicing physician, or the one who has lost his license to practice? Might not the standardized patient be able to support the physician in new learning in the way of the Bedside Clinics or, in some way, make it possible for the physician-in-trouble to relearn?

Epilogue

In looking back on any human endeavor, it is always interesting to see how diverse are

the motivations that shape that history. Altruism and egotism entwined, create the path, inspire the wisdom that has shaped the movement towards the way we are now teaching and testing the clinical skills of the young people who will be our future physicians. This is the standardized patient, a single educational innovation that had as much chance, or more, of not-being as being. It is the thread that held the inspiration until all was ready for the weaving—the golden winged rod entwined with oppositional energy that symbolizes the integration around which so much else has been explored and discovered. May that golden rod, now firmly planted, continue to inspire winged ideals in the midst of the inevitable conflict of opinions that will create the fertile soil for sustaining educational efforts as the search goes on for a better way to support the healers of today—and nurture those of tomorrow.



Notes

1. Howard S. Barrows, interview by author, Springfield, Ill., Feb. 18, 1996 (hereafter cited as Barrows interview).
2. Ibid.
3. "Hollywood Invades USC Medical School," *LA Herald-Examiner*, Sept. 27, 1965, A20.
4. "Models Who Imitate Patients: Paradise for Medical Students," *San Francisco Chronicle*, Sept. 28, 1965, 3.
5. Howard S. Barrows and Stephen Abrahamson, "The Programmed Patient: A Technique for Appraising Student Performance in Clinical Neurology," *Journal of Medical Education* 39 (1964): 802-5.
6. Barrows interview.
7. Gregory J. Magarian and Dennis J. Mazur, "Evaluation of Students in Medicine Clerkships," *Academic Medicine* 65 (1990): 341-45.
8. Barrows interview.
9. Ronald Harden, Mary Stevenson, W. Wilson

Downie, and G. M. Wilson, "Assessment of Clinical Competence Using Objective Structured Examination," *British Medical Journal* 1 (1975): 447-51.

10. Barrows interview.
11. Howard S. Barrows, "An Overview of the Uses of Standardized Patients for Teaching and Evaluating Clinical Skills," *Academic Medicine* 68 (1993): 446-51.
12. Barrows interview.
13. Ibid.
14. Ibid.
15. Ibid.
16. Robin M. Tamblyn and Howard S. Barrows, "Bedside Clinics in Neurology: An Alternate Format for a One Day Course in Continuing Medical Education," *JAMA* 243 (1980): 1448-50.
17. Barrows interview.
18. Ibid.
19. Ibid.
20. Arthur S. Elstein, Lee S. Shulman, and Sarah A. Sprafka, *Medical Problem Solving: An Analysis of Clinical Reasoning* (Cambridge: Harvard University Press, 1978).
21. Barrows interview.
22. A. Burri, K. McCaughan, and Howard S. Barrows, "The Feasibility of Using the Simulated Patient as a Means to Evaluate Clinical Competence of Practicing Physicians in a Community," in *Proceedings of the Fifteenth Annual Conference on Research in Medical Education* (Washington, D.C.: Association of American Medical Colleges, 1976), 295-99.
23. Ray Helfer and Joseph Hess, "An Experimental Model for Making Objective Measurements of Interviewing Skills," *Journal of Clinical Psychology* 26 (1970): 327-31.
24. Paula L. Stillman et al., "Construct Validation of the Arizona Clinical Interview Rating Scale," *Educational and Psychological Measurement* 37 (1977): 1031-38.
25. Paula L. Stillman, interview by author, Philadelphia, Mar. 16, 1996 (hereafter cited as Stillman interview).
26. Ibid.
27. Paula L. Stillman, Darrell L. Sabers, and Doris L. Redfield, "The Use of Paraprofessionals to Teach Interviewing Skills," *Pediatrics* 57 (1976): 769-74.
28. Stillman interview.
29. Robert M. Kretzschmar, "Evolution of the Gynecology Teaching Associate: An Education Specialist," *American Journal of Obstetrics and Gynecology* 131 (1978): 367-73.
30. Ibid., 368.
31. Stillman interview.
32. Ibid.

33. Ibid.
34. Paul Rutala and Paula L. Stillman, *The Non-Physician in Medical Education*, ed. Stillman (Tucson: University of Arizona, 1978), 26.
35. Stillman interview.
36. Ibid.
37. M. Angevine, interview by author, Tuscon, July 2, 1996.
38. Stillman interview.
39. Ibid.
40. Robin M. Tamblyn, "The Use of Standardized Patients in the Evaluation of Clinical Competence: The Evaluation of Selected Measurement Properties" (Ph.D. diss., McGill University, 1989).
41. Stillman interview.
42. Barrows, "An Overview of the Uses of Standardized Patients," 447.
43. Howard S. Barrows and M. J. Peters, eds., *How to Begin Reforming the Medical Curriculum* (Springfield: Southern Illinois University School of Medicine, 1984), i.
44. Barbara Gastel and David E. Rogers, eds., *Clinical Education and the Doctor of Tomorrow: Proceedings of the Josiah Macy, Jr. Foundation National Seminar on Medical Education, Adapting Clinical Medical Education to the Needs of Today and Tomorrow, Held June 15-18, 1988* (New York: New York Academy of Medicine, 1989), 112.
45. Thomas H. Meikle, interview by author, Philadelphia, Apr. 1, 1996 (hereafter cited as Meikle interview).
46. Ibid.
47. Linda J. Morrison and Howard S. Barrows, "Developing Consortia for Clinical Practice Examinations: The Macy Project," *Teaching and Learning in Medicine* 6 (1994): 23-27.
48. Paula L. Stillman and David B. Swanson, "Ensuring the Clinical Competence of Medical School Graduates Through Standardized Patients," *Archives of Internal Medicine* 147 (1987): 1049-52; Paula L. Stillman et al., "Assessing Clinical Skills of Residents with Standardized Patients," *Annals of Internal Medicine* 105 (1986): 762-71.
49. M. Brownell Anderson, Paula L. Stillman, and Youde Wang, "Growing Use of Standardized Patients in Teaching and Evaluation in Medical Education," *Teaching and Learning in Medicine* 6 (1994): 15-22.
50. Steven Muller, *Physicians for the Twenty-First Century: Report of the Panel on the General Professional Education of the Physician and College Preparation for Medicine* (Washington, D.C.: Association of American Medical Colleges, 1984).
51. Liaison Committee on Medical Education, *Functions and Structure of a Medical School: Accreditation and the Liaison Committee on Medical Education: Standards for Accreditation of Medical Education Programs Leading to the M.D. Degree* (Washington, D.C.: American Medical Association, 1991).
52. M. Brownell Anderson and Donald G. Kassebaum, eds., "Proceedings of the AAMC's Consensus Conference on the Use of Standardized Patients in the Teaching and Evaluation of Clinical Skills," *Academic Medicine* 68 (1993): 437-83; Terrill A. Mast and M. Brownell Anderson, eds., "Special Section: Annex to the Proceedings of the AAMC Consensus Conference on the Use of Standardized Patients in the Teaching and Evaluation of Clinical Skills," *Teaching and Learning in Medicine* 6 (1994): 2-35.
53. Anderson, Stillman, and Wang, "Growing Use of Standardized Patients," 15.
54. Harden, Stevenson, Downie, and Wilson, "Assessment of Clinical Competence."
55. Barrows, "Overview of the Uses of Standardized Patients."
56. Barrows interview.
57. Stephen Abrahamson, Minutes, Project Mousetrap Task Force Meeting, May 23-24, 1985, Philadelphia.
58. Ibid.
59. Reed G. Williams et al., "Direct, Standardized Assessment of Clinical Competence," *Medical Education* 21 (1987): 482-89; Nu Viet Vu et al., "Six Years of Comprehensive, Clinical, Performance-Based Assessment Using Standardized Patients at the Southern Illinois University School of Medicine," *Academic Medicine* 67 (1992): 43-50.
60. Cees P. M. van der Vleuten and David B. Swanson, "Assessment of Clinical Skills with Standardized Patients: State of the Art," *Teaching and Learning in Medicine* 2 (1990): 58-76.
61. Nu Viet Vu and Howard S. Barrows, "Use of Standardized Patients in Clinical Assessment: Recent Developments and Measurement Findings," *Educational Researcher* 23 (1994): 23-30.
62. Stillman et al., "Assessing Clinical Skills of Residents."
63. Paula L. Stillman, David B. Swanson, et al., "An Assessment of the Clinical Skills of Fourth-Year Students at Four New England Medical Schools," *Academic Medicine* 65 (1990): 320-26.
64. Alton I. Sutnick et al., "ECFMG Assessment of Clinical Competence of Graduates of Foreign Medical Schools: Educational Commission for Foreign Medical Graduates," *JAMA* 270 (1993): 1041-45; Alton I. Sutnick, Paula L. Stillman, and John J. Norcini, "Pilot Study of the Use of the ECFMG Clinical Competence Assessment to Provide Profiles of Clinical Competencies of Graduates of

- Foreign Medical Schools for Residency Directors," *Academic Medicine* 69 (1994): 65-67.
65. Daniel J. Klass, interview by author, Philadelphia, Mar. 14, 1996 (hereafter cited as Klass interview).
 66. *Ibid.*
 67. Barrows interview.
 68. Klass interview.
 69. Daniel J. Klass et al., "Portability of a Multiple Station, Performance-Based Assessment of Clinical Competence," in *Further Developments in Assessing Clinical Competence*, ed. Ian Hart and Ronald Harden (Montreal: Can-Heal, 1987), 434-42; Robin Tamblin et al., "How Standardized Are Standardized Patients?" in *Proceedings of the 27th Research in Medical Education Conference* (Washington, D.C.: Association of American Medical Colleges, 1988), 148-53; Robin M. Tamblin et al., "The Accuracy of Standardized Patient Presentation," *Medical Education* 25 (1991): 100-9.
 70. Klass interview.
 71. Meikle interview.
 72. Klass interview.
 73. Stephen Abrahamson, interview by author, Los Angeles, Mar. 11, 1996.
 74. Alton I. Sutnick, interview by author, Philadelphia, Mar. 14, 1996 (hereafter cited as Sutnick interview).
 75. Hadley L. Conn, Jr., "Assessing the Clinical Skills of Foreign Medical Graduates," *Journal of Medical Education* 61 (1986): 863-71; Hadley L. Conn, Jr., and Ronald P. Cody, "Results of the Second Clinical Skills Examination of the ECFMG," *Academic Medicine* 62 (1989): 448-53.
 76. Sutnick interview.
 77. *Ibid.*
 78. *Ibid.*
 79. Stillman interview.
 80. Sutnick et al., "ECFMG Assessment of Clinical Competence"; Sutnick, Stillman, and Norcini, "Pilot Study."
 81. Richard K. Reznick et al., "An Objective Structured Clinical Examination for the Licentiate: Report of the Pilot Project of the Medical Council of Canada," *Academic Medicine* 67 (1992): 487-93; Richard K. Reznick et al., "Large-Scale High-Stakes Testing with an OSCE: Report from the Medical Council of Canada," *Academic Medicine* 71 (1996): S19-S21.
 82. Educational Council for Foreign Medical Graduates, "Clinical Skills Assessment: Brief History," *Annual Report-Educational Commission for Foreign Medical Graduates* (Philadelphia: Educational Commission for Foreign Medical Graduates, 1996), 16-17.
 83. "Highlights of the 1995 Annual Meeting of the Board," *National Board Examiner* 42, no. 2 (1995): 1-3.
 84. "Ian R. Hart, MB, ChB Recipient of the 1996 Hubbard Award," *National Board Examiner* 43, no. 2 (1996): 1-3.
 85. Alton I. Sutnick, Miriam Friedman, and M. P. Wilson, "ECFMG International Ventures in Clinical Competence Assessment," in *Proceedings of the Sixth Ottawa Conference on Medical Education, Toronto, Ontario, June 26-29, 1994*, ed. Arthur I. Rothman and Robert Cohen (Toronto: University of Toronto Bookstore Custom Publishing, 1995), 311-12.

PEGGY WALLACE is Associate Adjunct Professor of Medicine and Director of Curricular Resources and Evaluation at the University of California, San Diego School of Medicine, where she is responsible for the teaching and assessment of clinical skills in the undergraduate medical school curriculum. She held a faculty position at the University of Southern California (USC) in the Department of Medical Education from 1977-1995 and was responsible for the re-introduction of standardized patients into the USC medical school curriculum starting in the mid-1980s. From 1993-1996, she guided the activities of the Southern California Consortium for the Assessment of Clinical Competency sponsored by the Josiah Macy, Jr. Foundation. She has served as a consultant to the National Board of Medical Examiners on the Standardized Patient Project and to various managed care organizations on physician-patient communication. She has also conducted workshops nationally and internationally for the World Health Organization on instructional technology, the use of video in medicine, and standardized patient training and case development. She is currently working on a book entitled *The Art and Practice of Using Standardized Patients in Clinical Education*, which will be published in 1998 as part of the Springer Medical Education series.

ACKNOWLEDGMENTS

The author wishes to express her thanks to the people who agreed to share their insights with her about their roles in the living history of this educational innovation: Stephen Abrahamson, Howard S. Barrows, Daniel J. Klass, Thomas H. Meikle, Paula L. Stillman, and Alton I. Sutnick. The author also wishes to thank Karen Garman and Diane Richards for their support and guidance as this manuscript came into being.