

## Using First-Year Students as Standardized Patients for an Objective Structured Clinical Exam for Third-Year Pharmacy Students<sup>1</sup>

Debra Sibbald

*Faculty of Pharmacy, University of Toronto, 19 Russell Street, Toronto, Ontario, Canada M5S 2S2*

Significant changes in the clinical context of health professional education and an impressive body of research on the measurement of clinical competence support the use of the Objective Structured Clinical Exam (OSCE) as the preferred means of performance-based assessment and the use of standardized patients as an adjunct to current teaching and evaluation methods. While there is research to support the use of standardized actors (SPs) as patients, using students as patients and raters has been little studied across disciplines and not in the profession of pharmacy. This work investigates the impact of using first-year pharmacy students as standardized patients. It examines the reliability, validity, feasibility and acceptability of using first-year, SP and faculty raters to evaluate performance in a third-year candidate self-medication OSCE. The major findings indicate that using freshman students is reliable and valid, is cost effective and provides learning benefits to participants.

### INTRODUCTION

At the Faculty of Pharmacy, University of Toronto, the author coordinates two sequential courses in non-prescription medications for large classes of undergraduate pharmacy students. A previous manuscript described the first three years of implementation and evolution including course design, teaching methodology, reinforcing and enabling strategies, case preparation, assessment tools, evaluations and examinations. Experiences were presented in managing issues and peer teaching, while fostering an interactive, motivating environment(1). One of the goals is to prepare the student to assume the role of a pharmacist who will accept accountability for community patient care by identifying, preventing and resolving problems relating to self-medication. This involves daily communication with patients, requiring accurate and concise evaluation of their global needs. A high level of skill in oral dialogue, interpersonal interactions and assessment is required to competently implement this professional responsibility. During these dialogues, the pharmacist is expected to efficiently process the analysis, synthesis and evaluation of relevant information, based on recall, understanding and application of cumulative knowledge. Whether that outcome has been achieved should be reflected in the type of evaluation at the conclusion of the two courses. In order to credibly evaluate each student's proficiency in the achievement of course objectives, a final, cumulative oral clinical skills examination was developed and implemented, using standardized patients. The development of this examination was reviewed in a second paper. (ACCP Innovations in Teaching Award, 1998)(2).

Since that time, the examination has undergone several years of modification in terms of student preparation and participation and in the use of first-year students as standardized patients. These improvements, described in this paper, help to bridge the gap between the classroom and practice and together with another project,(3) were the basis of the ACCP Innovations in Teaching Award, 2001.

### PROBLEM-BASED ASSESSMENT IN HEALTH CARE EDUCATION

PBA: Educational measurement has been enhanced in recent years through performance-based assessment (PBA). This was first described as evaluation methods that require "the examinee to demonstrate specific skills and competencies... to apply the skills and knowledge that have been mastered...the examinee's task is to construct an original response, which the examiner observes and evaluates"(5). Forms of PBA include four basic components: a reason for assessment, a specific performance to be evaluated, exercises which elicit that performance, and systematic rating procedures.

The popularity and acceptance of problem-based assessments in the arena of public education stems from their ability to measure the higher order critical thinking skills such as analysis, synthesis and evaluation, described in Bloom's taxonomy(6), and emphasize proficiency in processing skills: problem solving, comprehension, reasoning and meta-cognitive processes(7). This is represented schematically by this author in Figures 1 and 2.

The objective structured clinical examination (OSCE) has become the preferred form of PBA used in clinical competency assessment as it typically tests greater numbers of tasks, using greater numbers of examiners, thereby minimizing task and examiner related variance components.

**The OSCE:** The objective structured clinical examination (OSCE), first described by Harden et al.(8) has become an accepted format for evaluating clinical performance in the health care arena, as it is considered one of the most valid and reliable approaches, especially for a large group of candi-

<sup>1</sup> This article is based on a portfolio which was submitted to the ACCP Council of Faculties and presented during the innovation in Teaching Awards special session, ACCP Annual Meeting, Toronto, July 10, 2001. The title of the portfolio was "Innovative Enabling Strategies which Bridge the Gap from Learning to Practice."

*Am. J. Pharm. Educ.*, 65, 404-412(2001) received 8/9/01.

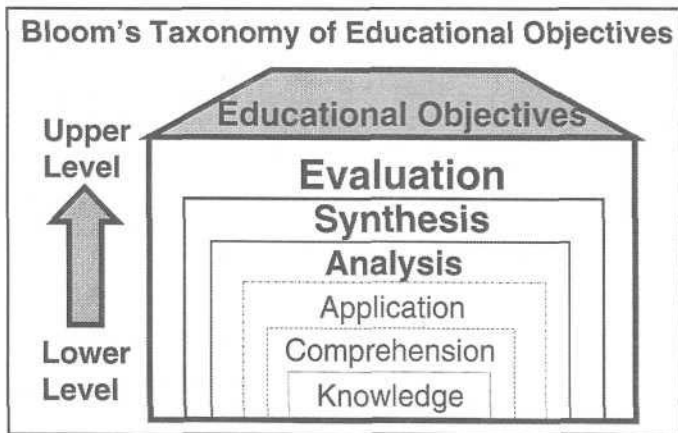


Fig. 1. Bloom's taxonomy of educational objectives.

dates(9). It is widely used in undergraduate curricula and for high stakes licensing and certification bodies in North America(10).

While OSCE structure may be varied, three key components are present: a highly structured format of task stations, objective scoring systems, and use of standardized patients (SPs) to portray the clinical problems in a consistent manner. Generally candidates must rotate through a series of timed stations chosen to reflect important components of clinical competence. The stations may require the examiner to use a checklist, which records behavior. Alternatively, global rating scales are used alone or with checklists to interpret clinical competence.

**Use of Standardized Patients:** For patient stations, either actual patients with stable physical findings, or standardized patients have been used(11). The standardized patient (SP), first proposed by Barrows and Abrahamson(12) is a normal individual who is taught to simulate every aspect of a patient's illness in a totally consistent manner so accurately that the simulation cannot be detected by a skilled clinician(13). Often, professional actors are used. The term 'standardized' patient has replaced the term 'simulated' patient because it underlines the major advantage of this technique: to provide a patient problem that will not vary from student to student. Standardized patients are trained to complete checklists and rating forms at the end of encounters.

**Raters:** An important consideration in test design is the determination of who should rate candidate performance. Adequate inter-rater agreement can be achieved through use of either SPs or faculty raters. Faculty raters are especially skilled at rating content-related tasks. SPs have been found to give more positive evaluations of communications skills of students than faculty, reflecting differences in perception in the role of observer and patient, and proposing that the patient, as the ultimate participant, is a more appropriate assessor of interpersonal skills(14).

#### RATIONALE FOR USING FIRST-YEAR STUDENTS AS PATIENTS AND RATERS

No study has examined the use of first-year students as standardized patients for senior student examinations. As subjects, first-year students who have not studied diagnosis or therapeutics, closely represent the unknowledgeable patient. As raters, such students, having taken a communication course, may have a heightened awareness of desirable communication skills. As learners, they may perceive a benefit in terms of application to

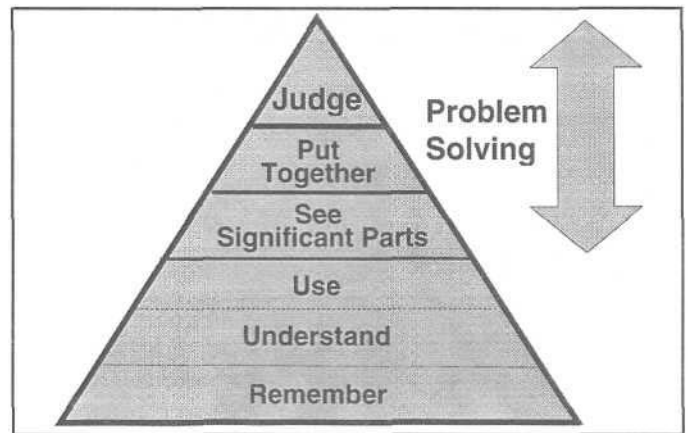


Fig. 2. Higher order critical thinking.

performance in senior years.

#### BACKGROUND INFORMATION AND EVOLUTION OF THE EXAM

The format of the examination evolved with changes over the first four years.

**Year One (1997):** In year one, feasibility was tested in a pilot project wherein each student had a single ten-minute interview with a simulated patient portrayed by a pharmacist. Both the role-playing pharmacist and an observing pharmacist rated the candidate. One practice was held prior to the exam between the candidate and a pharmacist teaching assistant (TA). Students requested a more realistic portrayal of the patient and an opportunity for more than one test interview.

**Year Two (1998):** In year two, each candidate had two interviews, and the patient role was played by a professional actor. The actor and the observing pharmacist rated the candidate. The practice format was changed to a small group interaction where the candidate experienced self, peer and instructor evaluation. The focus of the sessions was contemplation, discussion and rehearsal. Emphasis concentrated on behavioral skills, using the reflective student-centered approach espoused by Boud(15). Student feedback indicated the value and need for additional practices.

**Year Three (1999):** In year three, in addition to these group practices, thirty volunteer first-year first-year students were recruited to role-play a patient in an interview with a senior student, immediately following each class in second and third year. First-year students were selected in the hopes that their lack of therapeutic training would provide insights more closely parallel to the experience of the average consumer, and to increase their awareness and skills as future candidates. The overwhelmingly positive feedback from all students involved, who generated and unanimously signed a petition, was instrumental in allowing this initiative to be mandatory for future classes.

**Year Four (2000):** In year four, the first-year class of 120 students was divided into four cohorts. Two were assigned to the role-playing class practices with the second and third year students. A third cohort conducted individual private practices with a candidate (SR) prior to the exam, where the focus was feedback from the first-year students about organization and communication, rather than content. The fourth cohort underwent standardized training to achieve consistency in both role-portrayal and use of the assessment tool. They were then added to the oral examination patient pool, to enable each candidate

**Table I. Background Evolution of Oral Examination (OSCE)**

Year	Number of tests	Patients	Raters	Individual practices	Class practices
1 (1997)	1	TA	2TAs	SR with TA	
2 (1998)	2	SP	SP + TA	1 group (TA with 6SR)	
3 (1999)	2	SP	SP + TA	1 group (TA with 6SR)	
4 (2000)	4	2 SPs + 2 JRs	SP+TA  JR+TA	1 group (TA with 6 SR)  + 1 private (JR with SR)	16 in 3rd yr (JR with SR)  + 15 in 2nd yr (JR with SR)

to have four interviews with a standardized patient: two with a first-year student and two with a professional actor. This background evolution of the structure of the oral examination and the formative practices led to the innovative OSCE which is the basis of this paper (see Table I).

## PROJECT DEVELOPMENT

### Purpose of Innovation

The purpose of this innovation is to use first-year first-year students as standardized patients in the third year self-medication OSCE at the University of Toronto, Faculty of Pharmacy. The impact of this work has been investigated by examining the psychometric properties of reliability and validity of using first-year student, SP and faculty raters to evaluate performance, and through qualitative surveys of candidates and first-year students, over two years of implementation (2000, 2001)

### Project Questions

Psychometric Questions:

- Reliability: Using first-years students vs SPs as patients,
- Is there an effect on the faculty, shown by the scores they generate?
  - Is there an effect on candidate performance, shown by the scores achieved?

Validity: Using first-years students vs SPs as patients,

- Is there a difference in the scores generated by patient raters vs TA raters?
- Is the OSCE a valid measure compared with other assessments of performance?

Qualitative Questions:

- How do the exam candidates and first-year students who participated regard the project as a learning experience?

## PROJECT DESIGN

### Format of the OSCE

The OSCE is designed to measure competencies acquired by candidates during completion of two sequential self-medication courses taught during second and third year. The cumulative exit OSCE occurs in March, 10 weeks after completion of the second course. The examination is administered over two successive days and tests approximately 120 candidates. Each day, the candidate completes two ten-minute interviews

followed by a five-minute post encounter assessment period. There is an interval between stations of 45 to 60 minutes. To facilitate the process, five parallel circuits are administered in three consecutive sessions on each day. The first day and second day interviews are scheduled for each candidate at the same times on each day, so that a 24-hour time lapse occurred for each. Table II illustrates the OSCE circuits for the first day of the OSCE.

The first-year of the project (2000), a standardized actor (SP) from the University of Toronto Standardized Patient Program played the patient on day one and a first-year pharmacy student played the patient on day two. This sequencing was reversed the following year, in response to student feedback. The performance of candidates at each station is observed and rated by a practicing pharmacist (TA) who is a teaching assistant or professor at the Faculty. Table III outlines the patients and raters for the OSCE for March, 2000.

### Quality Control of Consistency in Role Delivery

For each examination day, all patients and TA raters who participate in the same case watch a preliminary enactment of the role to establish consistency in role-delivery. The course instructor rotates through each set of five rooms assigned per case to observe and comment on consistency in role-delivery and consistency in use of the assessment tool.

### Description of Data Collection

During the encounter, the TA observer completes a case-specific five item task-specific checklist, which relates to problem identification and treatment. Following the encounter, the candidate leaves the room. The TA, the patient and the candidate, complete a global rating scale, which is the same for all stations. Patient and TA raters are carefully cautioned to resist the inclination to compare impressions. Their scores are included in data analysis.

Candidate self-assessments are not included in data analysis. They are carried out in order to give the candidate an opportunity to reflect upon and evaluate their performance after each interview. It is an opportunity to garner insights in preparation for the next test interview. At the completion of the four-station OSCE, each candidate completes a survey to collect data regarding their impressions of the OSCE.

### Method for Recording Incidents

Each interview is audio-taped by the TA observer. They

**Table II. The OSCE circuits: Day one**

Time	Activities for Day One									
7:45-8:15	<b>Briefing</b>									
8:15-8:30	<b>Session One</b>									
	<b>Practice run of case 1 or 2 with 1 simulated patient and 1 teaching assistant</b>									
	<b>Case 1</b>					<b>Case 2</b>				
<b>Room</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>	<b>I</b>	<b>J</b>
<b>Candidate number</b>										
8:30-8:45	1	2	3	4	5	21	22	23	24	25
8:45-9:00	6	7	8	9	10	26	27	28	29	30
9:00-9:15	11	12	13	14	15	31	32	33	34	35
9:15-9:30	16	17	18	19	20	36	37	38	39	40
<b>Switch cases</b>										
9:30-9:45	21	22	23	24	25	1	2	3	4	5
9:45-10:00	26	27	28	29	30	6	7	8	9	10
10:00-10:15	31	32	33	34	35	11	12	13	14	15
10:15-10:30	36	37	38	39	40	16	17	18	19	20
10:30-10:45	Break									
10:45-11:00	<b>Session Two</b>									
	<b>Practice run of case 3 or 4 with 1 simulated patient and 1 teaching assistant</b>									
	<b>Case 3</b>					<b>Case 4</b>				
11:00-11:15	41	42	43	44	45	61	62	63	64	65
11:15-11:30	46	47	48	49	50	66	67	68	69	70
11:30-11:45	51	52	53	54	55	71	72	73	74	75
11:45-12:00	56	57	58	59	60	76	77	78	79	80
<b>Switch cases</b>										
12:00-12:15	61	62	63	64	65	41	42	43	44	45
12:15-12:30	66	67	68	69	70	46	47	48	49	50
12:30-12:45	71	72	73	74	75	51	52	53	54	55
12:45-1:00	76	77	78	79	80	56	57	58	59	60
1:00-1:45	<b>Lunch</b>									
1:45-2:00	<b>Session Three</b>									
	<b>Practice run of case 5 or 6 with 1 simulated patient and 1 teaching assistant</b>									
	<b>Case 5</b>					<b>Case 6</b>				
2:00-2:15	81	82	83			96	97	98		
2:15-2:30	84	85	86			99	100	101		
2:30-2:45	87	88	89			102	103	104		
2:45-3:00	90	91	92			105	106	107		
3:00-3:15	93	94	95			108				
<b>Switch cases</b>										
3:15-3:30	96	97	98			81	82	83		
3:30-3:45	99	100	101			84	85	86		
3:45-4:00	102	103	104			87	88	89		
4:00-4:15	105	106	107			90	91	92		
4:15-4:30	108					93	94	95		
4:30-5:00	<b>De-briefing</b>									

**Table III. Patients and raters for the OSCE**

Stations	Raters	
Day One: Standard actors (SPs) as patients		
Station 1	SP rater # 1	TA rater # 1
Station 2	SP rater # 2	TA rater # 2
Day Two: Juniors (JRs) as		
Station 1	JR rater#1	TA rater # 1
Station 2	JR rater # 2	TA rater #2

note on the schedule any untoward events, the candidate involved and the tape number and time for future review and scrutiny. In their self-assessments, the candidates also have the

opportunity to record any untoward events and their impressions of the incident for future evaluation.

**Case Design**

A total of twelve cases are written for the OSCE. Six are used each day in pairs over three administered sessions. Each sessional pairing is designed to be similar in terms of complexity of the content, and complexity of the patient. However, to ensure content validity, the clinical scenarios present as wide a range of situations as a self-medication advisor would be expected to handle. The sampling is twelve of thirty curriculum topics or 40 percent. Each case is an interaction between a simulated patient and the candidate, who assumes the role of a practicing pharmacist. The cases are designed to address the course

objectives in terms of content, processing the delivery of pharmaceutical care, and interpersonal skills, in addition to assisting patients with special needs. Cases are structured to test performance in four separate areas of professional competency which the candidate is expected to identify and solve. These include:

- a communication challenge;
- a special need, such as a social, emotional or cognitive issue;
- a key component in history-taking relating to medical conditions, family history, allergies or medications;
- the identification, prevention or resolution one or two drug-related problems.

These key points are summarized in a short, five point checklist of task-specific items that is required of the candidates relating to problem identification and treatment. Integration of this knowledge with skills is required to pass the interview and is reflected in the scoring of the assessment tool in the domain that addresses overall impression (knowledge and skills).

The candidate is expected to dialogue with the patient effectively. The scope of the material is cumulative over the two courses. The cases are reflective of realistic patients in pharmacy practice. This content validity is assured since they are real patient situations encountered by the course instructor in her pharmacy practice as an advisor for self-medication. In that practice site, the instructor has no other tasks except dialoguing with patients in the over-the-counter (OTC) sections in a community pharmacy and typically will advise 120-130 patients during each shift. The cases are written by the case instructor and reviewed by another instructor for content and level of difficulty.

### **Description of the OSCE Practices**

Candidates are provided with extensive information about the purpose and format of the testing in advance of the OSCE. This is to prevent candidate misconceptions about level and quality of performance expected by the test design, and to allow formative feedback to enhance performance skills. At the commencement of the second year course, a training session introduces students to the global rating scale. The purpose, use and interpretation are outlined using videotapes of student interviews.

Three types of formative practices are experienced by candidates prior to the OSCE. First, after each hour case-based class on a self-medication topic in both second and third year, they observe or participate as a pharmacist in ten-minute practice interviews conducted in front of the entire class. The patient is a first-year student and the case, an applied version of the same topic, is written by a class team. The global rating scale is scored by the first-year student, the pharmacist and the team who writes the case. Feedback comments are elicited from each of the raters and a class discussion is generated. Secondly, a formal practice of ten-minute interviews is conducted as part of the third-year Pharmacy Practice laboratory, in teams of six students with a TA. The students role-play patients for each other in turn and the cases are written by the course instructor for self-medication. The global rating scale is used by each observer which generates discussion and suggestions for improvement. Lastly, each candidate has a private ten-minute practice interview with a first-year student prior to the OSCE. The cases are written by the self-medication course instructor. The first-year student rates the candidate using the

global rating scale and gives immediate oral feedback regarding their feelings as a patient during the dialogue; and the organization, focus, and communication skills of the candidate. Each first-year student repeats the role-portrayal with two or three candidates.

### **Description of the First-Year Students**

The first-year students are recruited from a pharmacy social science course, which examines the perspective of the patient. These students complete a full semester course in communication skills. Sixty students who score highest on their entrance to pharmacy test of oral and written communications skills are selected for participation in the exam. A total of six patient roles are delivered by first-year students. They receive their roles, but no description of content checklist items, ten days in advance. First-year students are then trained by the self-medication instructor both for consistency in using the global rating scale, using videotapes of student interviews, and in standardizing their roles. Each first-year student repeats the assigned role for no more than four ten-minute interviews on the examination day, in order to utilize all sixty students, and to protect against fatigue in these inexperienced role-players.

The remaining students of the first-year class are sorted into sections for participation in either the class practices (one per student) or the private individual interviews (two to three per student). Participation in the OSCE or the practices are part of their curriculum cross-course requirements. After each experience, first-year students complete a survey to collect data regarding their impressions.

### **Confidentiality Issues**

Prior to the examination, each candidate and each first-year student are required to read the code of conduct governing test security for examinations at the University. They sign a confidentiality agreement, which prohibits discussion of the nature of the patient roles or stations until the examination is completed. A literature review of test security using SPs reveals little or no problem with test security, especially when the SP application is measuring process rather than the ability to recall content and produce answers. Security should be monitored, but the issues should not restrict use(16). It is hoped that operationally, an element of uncertainty among the candidates is present as to the identity of the cases, and the use of 12 different cases and stringent efforts to enforce security will discourage any dissemination of confidential checklist items by first-year students.

During the examination, candidates are sequestered in separate rooms with invigilators for each parallel stream to ensure there is no verbal contact. Incoming candidates are sequestered before outgoing candidates leave the building. Each candidate and first-year student are required to disclose in advance, any close relationship with a student in the other year, and the pairings are scheduled to avoid combining first-year students and candidates who may know one another.

### **CHALLENGES IN IMPLEMENTATION**

There are issues which require more careful planning in using of students vs actors. Although students are accessible, careful scheduling around their timetable is needed to make training available at times which are convenient and reasonable in duration, yet close enough in proximity to the examination date to afford a measure of test security. In consideration of their workload, the investigator is careful to schedule the examina-

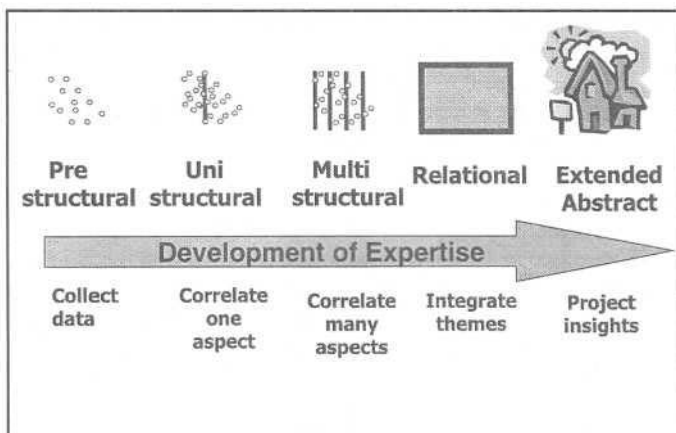


Fig. 3. Author's schematic representation of the SOLO taxonomy of Collis and Biggs.

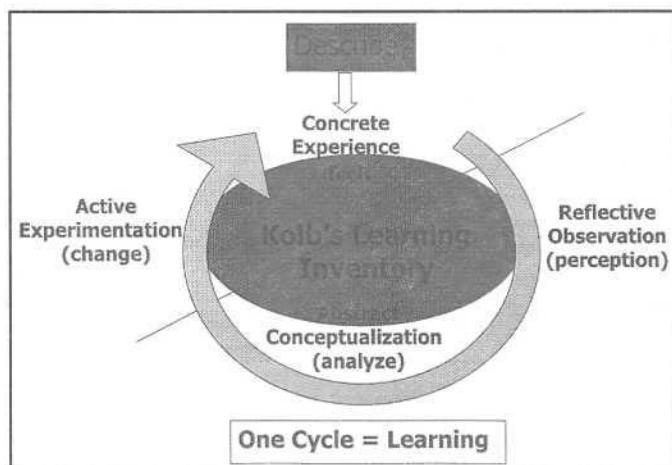


Fig. 4. Feedback based on learning cycle.

tion date on a day which does not conflict too significantly with workload or examinations in other first-year courses, and this is a weekend day. Students are not reimbursed for their participation and have to provide their own transportation which is problematic for some since the OSCE is not held during regular classes. Finally, in order to affect a high degree of realism in role-delivery, all first-year student cases are written to portray patients of this age group, whereas greater age variability is possible with SPs. However, since the first-year student cohort is multi-cultural, it is possible to direct a variety of ethnicities into their roles, and this is not possible with the SPs, who are a much smaller cohort.

### MODIFICATIONS

Initially, SPs were used on day one (Saturday) and first-years students were used on day two (Sunday). Feedback from first-year student was useful in suggesting reversing this order, giving them an opportunity to use Sunday for relaxing, work or study for the coming week. Candidates made a similar suggestion, since they reported being more at ease with the first-year students than the SPs during the interviews, and felt it would lessen their anxiety for SP interviews if they began with first-year. This suggestion made scheduling easier for the instructor and was implemented the following year.

### ASSESSMENT

#### Description of Assessment Tool: a Global Rating Scale

The assessment tool used to gather data for analysis is a global rating scale, which is the same for all stations and used by all the raters for the exam. It was designed by a rhetorician in consultation with clinical educators from several health care disciplines to test processing skills which reflect competence. The scale consists of four domains, which assess empathy, coherence (organization/focus), verbal skills and non-verbal skills on a five-point scale. A fifth domain is the overall impression of the integration of knowledge and skills in the performance, also on a five-point scale.

The five-point scale can be interpreted in terms of the five levels of learning described by Collis and Biggs as the Structure of the Observed Learning Outcome (SOLO) taxonomy(17). This scheme is well-designed to reflect the quality of learning of a complex set of skills such as process analysis and discriminates between performance levels described as pre-structural (novice), uni-structural (beginner), multi-structural (competent), relational (proficient) and extended abstract

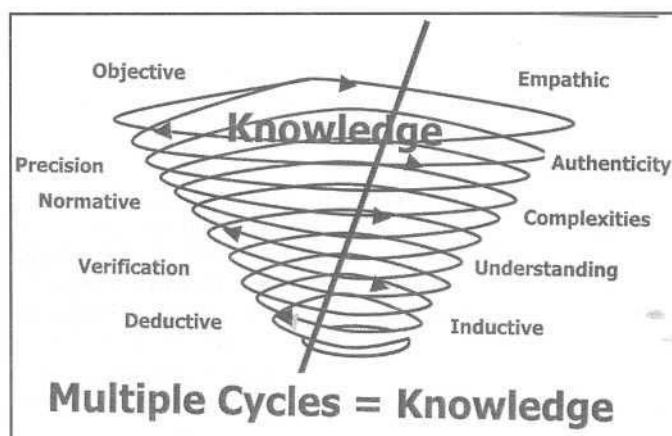


Fig. 5. Iterative formative practice.

(expert or insightful). This author has represented the hierarchy schematically in Figure 3.

Formal evaluation has shown that this instrument has quite acceptable psychometric properties(18) and rewards efficiency and mastery rather than thoroughness. Appendix C is the global rating scale used for the OSCE.

#### Description of Student Feedback

Student perceptions (both candidate and first-year student) of the examination project are determined by surveys. Quantitative survey questions are graded using a Likert scale with the response "Strongly Agree" = 5 and the response "Strongly Disagree" = 1. Students are asked to complete this survey anonymously. All first-year students are also asked to complete a free response feedback survey after participating in either the practices or the examination. The author designed this form to capture the learning cycles described by Kolb(19). The four-part form directs the student to describe their experience, reflect on their feelings during the encounter, analyze why this was so, and suggest modifications for future administrations to improve the program. In this way, each iterative experience expands the cycle of learning about the project into a widening cone of knowledge that increases the breadth and depth of understanding. This knowledge is a balanced assessment, since it is based on both deductive or objective and inductive or empathic reasoning. See Figures 4,5 and 6.

Summary of Cooperative Experience	Name:
Recall your experience:	
What are you feeling about?	
Analyze (why things happened this way, why you felt this way)	
Suggestions for further action (based on the reflections above)	

Fig. 6. Feedback form for first-year student participation.

## EVALUATION AND RESULTS

### Psychometric Results

**First-year of Innovation (2000):** One hundred and eight candidates were examined. The overall examination score was 73.1 percent with a standard deviation of 6.7 and a range from 53.3 percent to 86.5 percent. Four students failed (3.7 percent) and 15 (14 percent) received honor-status marks.

**Second year of Innovation (2001):** One hundred and eighteen candidates were examined. The overall examination score was 75.14 percent with a standard deviation of 5.9 and a range from 58.8 percent to 87.7 percent. One student failed (0.1 percent) and 32 (27 percent) received honor-status marks. Statistical tests are reported for results of the exam for 2000. Statistical tests for 2001 are still in progress, in order to compare correlations with future performances and establish predictive validity.

### Reliability: Influence on Faculty Scores

A paired *t*-test comparing means was done to examine if the use of 1st years' vs SPs as patients had a systematic influence on scores by faculty raters (TAs). Data analysis showed there were no significant differences in the TA scores for station 1 and station 2 ( $t_{104} = 1.84, P = 0.07$ ) when the SPs played the patients ( $\chi = 4.26 \pm 1.07$ ) vs. when first year student played the patient, ( $\chi = 4.49 \pm 0.96$ ), with the difference in means amounting to less than 0.23 standard deviation size.

### Reliability: Influence on Candidate Scores

A paired *t*-test comparing means was done to examine if the use of SPs vs. first year students as patients had a systematic influence on candidates scores. Data analysis showed there were no significant differences ( $t_{104} = 0.27, P = 0.79$ ) in the mean SP scores ( $\chi = 4.93 \pm 0.88$ ) and the mean JR scores ( $\chi = 4.90 \pm 0.83$ ), with the difference in means equaling approximately 0.05 standard deviations.

However, it is worth noting that the means of both SP ( $\chi = 4.93 \pm 0.88$ ) and JR ( $\chi = 4.90 \pm 0.83$ ) raters were higher than the means of TA raters, (SP patients,  $\chi = 4.26 \pm 1.07$ ; JR patients,  $\chi = 4.49 \pm 0.96$ ) and to the same level. Thus both types of patient raters marked equally more generously as compared to TA raters.

### Concurrent Validity: Comparison of TA Scores with SP and First-Year Scores

Finally, the correlations between SP vs. TA scores and first-year student vs. TA scores were assessed for station 1 and 2. Strong Pearson correlations verified that the TA and first-

year student scores were statistically similar for each station as were the TA and SP scores. The Pearson correlations for the SP vs. TA scores were 0.378 ( $P = 0.0001$ ) for station 1 and 0.579 ( $P = 0.0001$ ) for station 2. The Pearson correlations for the first-year student vs. TA were 0.584 ( $P = 0.0001$ ) for station 1 and 0.619 ( $P = 0.0001$ ) for station 2. The correlations for the averaged scores for TA vs. SP raters was 0.565 ( $P = 0.0001$ ) and for TA vs. JR raters was 0.646 ( $P = 0.0001$ ).

These strong Pearson correlations verified that the TA and JR scores were statistically similar for each station as were the TA and SP scores. Therefore, SP and JR raters produced correlations that were both reasonable and the same compared to TA raters. These results indicate using either SPs or first-year students as raters is valid in comparison to professional TAs, when scoring with a global rating tool.

### Concurrent and Predictive Validity: Correlations with Other Measures of Performance

Data showed the OSCE produced significant correlations with other measures of performance, including the two written exam in the same course ( $P = 0.0001$ ), a question and answer format oral exam in therapeutics ( $P = 0.011$ ), averaged traditional science courses for third year ( $P = 0.001$ ), GPAs from current ( $P = 0.002$ ) and previous ( $P = 0.0001$ ) years, and the averaged results of two clinical experience rotations (community and hospital) held the following year as an exit requirement ( $P = 0.017$ ). These results indicate that using either SPs or first-year student as patients and raters produce valid results with concurrent and prior performance measures and show predictive validity for future performance

### Student Feedback Results

In general, the replies to the survey over two years probed student perceptions of learning (Table IV). Survey results recorded candidates' views of the value of the project in terms of enhancing learning through exam preparation, measuring readiness to practice, usefulness of practices, value of use of first-year students, exam organization and support of continuation of the exam and this format. Free response feedback from the candidates reflected their positive feelings about the qualitative value of this experience. They commented on the increased gains in the professional practice, clinical problem-solving and communication skills. They were unanimous in their support of the continuation of this program as a facilitative way to acquire professional competencies.

Free response from first-year students were overwhelmingly positive in assessing the value of this project (Table V). First-year students reported gains in many areas: ability to experience peer-teaching, ability to network with senior students, training for their own future OSCE, increased tolerance, a balanced health care perspective in terms of the patient's views and needs, increased communication skills, and introduction to the long-term goals of the program. They felt empowered by the ability to effect program development through the use of formative feedback.

## DISCUSSION

Psychometric testing show that using first-year students as patients and raters produces results that are reliable and valid compared to previous, concurrent and predictive measures of performance. It is also cost effective. While operational costs for invigilators and faculty raters are the same on either day, there are no additional expenses for using first-year students as

**Table IV. Candidate evaluation survey of project for 2000 2001**

	Agree strongly	Percent
General:	/5	
My preparation for the exam enhanced my learning.	4.4	88
I was satisfied with my performance in the oral exam.	3.8	76
The oral exam is a good measure of my readiness to practice OTC skills.	4	80
Oral exams in this course should be continued.	4.3	86
Class Practices		
Writing a practice case reinforces the objectives of the oral exam.	4	80
Watching class practices clarified my understanding of the exam format	4	80
The class practice orals should be continued in both 2nd and 3rd year.	4.1	82
Individual Practices		
I received useful feedback from the 1st year student during my practice.	3.2	64
First-year students should not be part of practice orals: A classmate can play the patient.	5	30
The individual practices with the 1st year students should be continued.	4	80
Lab Practices		
My practice oral in the Pharmacy 329 lab helped prepare for the oral exam.	4.3	86
I received feedback, which clarified my strengths/areas for improvement.	4	80
The Pharmacy 329 lab practice orals should be continued.	5	100
Use of First-year Students as Standardized Patients		
The examination should involve only actors, not 1st year students.	1.1	22
I prefer having both actors and 1st year students as patients in the exam.	3.7	74
The use of 1st year students as patients reinforced learning for 3rd year students.	4	80
The use of 1st year students for the examination should be continued.	4	80
Organization		
The oral exam was well organized.	4.1	82
I prefer that no talking be allowed when waiting in the common room.	2.9	58
It is not necessary to audio tape the exams.	3.2	64
It was fair to be assessed by both the actor and the pharmacist observer.	4.1	82
My performance can be sufficiently assessed in 2 interviews.	3.1	62
I prefer 2 interviews if 4 interviews are only possible using 1st year students.	1.2	24
I would prefer 4 interviews regardless of who plays the patient.	4	80
Free Responses		
It increased confidence in ability to deliver pharmaceutical care and solve clinical problems (88%)		
It increased skills in oral communication (90%)		
It increased skills in professional practice (88%)		

**Table V. Comments of first-year student free response to project for 2000 and 2001**

• ability to experience peer-teaching (90%)
• ability to network with senior students (88%)
• training for their own future OSCE (95%)
• increased tolerance (64%)
• a balanced health care perspective in terms of the patient's views and needs, (78%)
• increased communication skills (79%)
• introduction to the long-term goals of the program (90%)
• empowered by the ability to effect program development through the use of formative feedback. (66%)
• impressed with breadth/depth of candidate's knowledge (78%)
• could see differences between candidates in skills level (knowledge) (80%)
• could see differences between candidates in ability to relate to patient (communication) (86%)

patients, whereas the cost for using professional actors ranges from \$150 to \$200 per student, and include financing of SPs, and a trainer(16). However, it is also apparent from the feedback from candidate and first-year student participants that they value this experience as an enhancement to learning. The surveys produce a long list of perceived enrichments, making this innovation a two-edged sword in serving learning needs of both exam candidates and first-year students. It is important to examine the reasons this occurs.

**Benefit to Candidates:** It is useful to reflect upon where the use of first-year students as standardized patients fits into the continuum of teaching and evaluation by considering what the first-year participant is able to do. The first is demonstration: first-year students can be used as standardized patient instruc-

tors to demonstrate something which is difficult or uncomfortable to learn from real patients. Beyond that, first-year standardized patients can be employed in the summative evaluation of learning, to judge whether educational goals have been achieved. Because of their judgmental character, the psychometric properties of validity and reliability are of central importance here. Finally, first-year students can facilitate learning through formative evaluation exclusively for the purpose of feedback. Candidates who undergo training practices with first-year students can be assured of a non-judgmental assessment that allows them to reveal what they do not know for the purposes of learning without fear they will be judged unfavorably. For this purpose, the psychometric qualities of the test are relatively unimportant.

The behaviors a pharmacist must learn are an integration



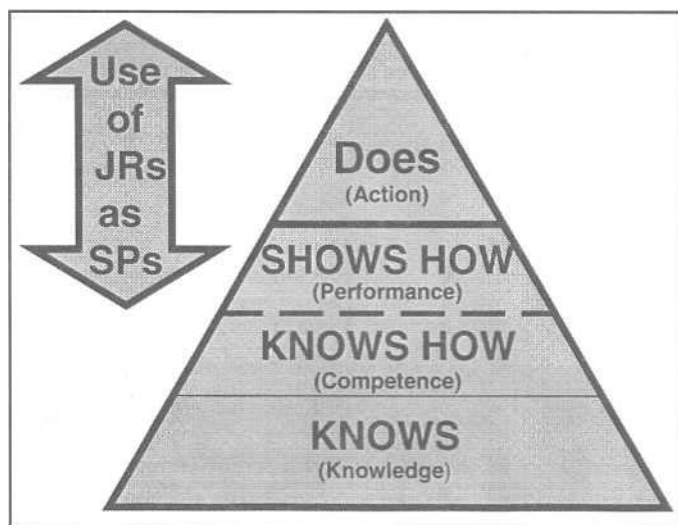


Fig. 7. Pyramid of professional behaviors.

of a complex set of components illustrated in the diagram shown in Figure 4. The well-informed professional must know how to use the knowledge he possesses. Nevertheless, readiness for practice is determined by the ability to demonstrate that he is not only skillful at processing this knowledge during observation and assessment but can carry out appropriate action independently. It is at the upper two levels of this pyramid that first-year students are most useful as standardized patients. Both the OSCE itself and the formative practices are arenas in which the first-year student can enhance the learning of the candidate through feedback and assessment. The private individual practices between first-year and senior students before the OSCE are a way to enable autonomous action in the absence of any observation.

**Benefit to First-Year Students:** The first-year students are encouraged to recognize that their subjectivity is engaged purposefully, and to use this lens to inquire into perspectives and interpretations, to shape new questions through examining assumptions, and to explore feelings in relationship to learning. Through the direction of the feedback form, they acquire a reflective practice that encourages them to acknowledge, embrace, challenge or augment gains in scholarship.

#### IMPLICATION FOR FACULTY WORKLOAD AND FINANCES

Operational factors such as organization and training may be one area in which financial and time constraints are at odds. It is far simpler in terms of faculty time management to share or delegate the training and scheduling demands of a small group of thirty professional actors with a professional trainer, if the financial budget supports this expenditure. In contrast, there are considerably increased demands on instructor time which is not assisted in any way, to deliver the same outcomes for the sixty students involved in the OSCE, including sorting the class into the group selected for the exam, and training for consistency in role-portrayal and use of the rating tool. The use of the balance of first-year students for the practices also requires the instructor's time for training and scheduling.

#### CONCLUSION

This innovation establishes the value and role of networking between first-year and senior students in education and evalua-

tion. Reliable results and valid correlations with other measures of performance make this a model for teaching that is easily feasible not only in other schools of pharmacy but in the education of other health professionals. This is a method, transferable to other providers, that reduces costs; generates a database of valid, consistent, reviewed and successful cases; and provides an experiential venue that is more practical than outside placements in bridging the gap between learning and readiness to practice. It forges a path that may lead to enhanced outcomes for both health care professionals and the patients they serve.

**Acknowledgements.** The author gratefully acknowledges the assistance and support of the following people in contributing to this project: Dr. Cleo Boyd, who developed the global rating scale used as the measurement instrument in the study; Dr. Heather Boon, professor of the first-year pharmacy social administration course, for enthusiastic collegial support and prompt and organized assistance in the sorting of first-year students; and the classes of first-year students for their participation. In willingly giving their time, energy, enthusiasm and insights to this project, they too learned from this experience.

#### References

- (1) Sibbald, D.J., "Innovative, problem-based, pharmaceutical care courses for self-medication," *Am. J. Pharm. Educ.*, **62**, 109-118(1998)
- (2) Sibbald, D.J., "Oral clinical skills examination: An innovative reinforcing strategy for nonprescription medication courses," *ibid.*, **62**, 458-463 (1998)
- (3) Sibbald, D.J., "Bridging the gap from classroom to practice: PBL Students develop consumer website for nonprescription drugs," *ibid.*, **64**, (4) 339-348 (2000)
- (4) Boud, D., Keough, R., and Walker, D., *Reflection: Turning Experience into Learning*. Kogan Page, London, (1985)
- (5) Stiggins, R.J., "Design and development of performance assessments," *Educational Measurement: Issues and Practice*, Fall, 33-41(1987)
- (6) Bloom, B.S. et al., *Handbook on Formative and Summative Evaluation of Student Learning*, McGraw-Hill Book Company, New York NY (1971) p.932.
- (7) Linn, R.L., Baker, E.L. and Dunbar, S.B., "Complex, performance-based assessment: Expectations and validation criteria," *Educational Researcher*, **20**(8), 15-21(1991)
- (8) Harden, R.M., Stevenson, M., Downie, W.W. and Wilson, G.M., "Assessment of clinical competence using an objective structured clinical examination," *Brit. Med. J.* **1**, 289-296(1975)
- (9) Harden, R.M., "Assess clinical competence - An overview," *Medical Teacher* **1**, 289-296(1979)
- (10) Van der Vleuten, C.P.M. and Swanson, D.B., "Assessment of clinical skills with standardized patients: State of the art," *Teaching Learning Med.* **2**, 58-76(1990).
- (11) Davis Feikert, J.A., Harris, I.B., Anderson, D.C., Bland, C.J., Allen, S., Poland, G.A. and Satran, L., Miller, W.J., "Senior medical students as simulated patients in an objective structured clinical examination: Motivation and benefits," *Medical Teacher* **14**(2/3), 167-177(1992)
- (12) Barrows, H.S., "Simulated (standardized) patients and other human simulations," Health Science Consortium, Chapel Hill NC (1971)
- (13) Barrows, H.S. and Abrahamson, S., "The programmed patient: A technique for appraising student performance in clinical neurology," *J. Med. Educ.*, **39**, 802-805(1964)
- (14) Cooper, C. and Mira, M., "Who should assess medical students' communication skills: Their academic teachers or their patients?" *Med. Educ.*, **32**, 419-421(1998)
- (15) Boud, D., Keough, R. and Walker, D., *Reflection: Turning Experience into Learning*. Kogan Page, London, (1985)
- (16) Stillman, P.L., "Technical issues: Logistics," *Acad. Med.*, **68**, 464-476(1993)
- (17) Collis K.F. and Biggs J.B., "Classroom examples of cognitive development phenomena: The SOLO taxonomy" Australia (1979)
- (18) Hodges, B., Regehr, G., Hanson, M. and McNaughton, N., "An objective structured clinical examination for evaluating psychiatric clinical clerks," *Acad. Med.*, **72**, 715-721(1997)
- (20) Kolb, D.A., Individual learning styles and the learning process (Report 535-571), Massachusetts Institute of Technology., Cambridge MA.