Pharmacy Education (2023) 23(1) 314 - 320
https://doi.org/10.46542/pe.2023.231.314320

RESEARCH ARTICLE

Comparison of pharmacy student, evaluator, and standardised patient assessment of Objective Structured Clinical Examination (OSCE) performance

Ashleigh L. Barrickman[1], Marina Gálvez-Peralta[2], Lena Maynor[1]

[1] Department of Clinical Pharmacy, West Virginia University School of Pharmacy, Morgantown, West Virginia, United States
[2] Pharmaceutical Science, West Virginia University School of Pharmacy, Morgantown, West Virginia, United States

Keywords
Assessment
Objective Structured Clinical Examination
OSCE

Correspondence
Ashleigh L. Barrickman
Department of Clinical Pharmacy
West Virginia University School of Pharmacy
Morgantown
West Virginia
United States
alandis1@hsc.wvu.edu

Abstract

Objective: The study aimed to (1) determine if standardised patient (SP) assessment of student communication differs from faculty assessment and (2) identify factors that may affect their assessment scoring since differences between evaluator scoring could impact students' learning and feedback. **Methods:** Students completed four OSCEs in spring 2022. Communication skills were scored using a global assessment rubric completed by a faculty evaluator, SP, and student. SPs and students rated their confidence in student recommendations on a scale of 1-5 (1=lowest; 5=highest). Spearman correlation and Wilcoxon matched-pairs signed rank tests were used for analysis. **Results:** SP global assessment scores were statistically significantly higher than faculty scores. SP global assessment and confidence scores were strongly correlated. **Conclusion:** Differences in OSCE scoring among evaluator types and factors that may impact scoring are important to explore. SP scores were consistently higher than faculty scores, which should be considered when using OSCEs as assessments.

Introduction

The use of objective structured clinical examinations (OSCEs) in health professions education is widely accepted as the gold standard for assessing clinical and communication skills (Shirwaikar, 2015; Urteaga *et al.,* 2015; Weaver *et al.,* 2022; Kristina, Wijoyo, 2018). Approximately 77% of pharmacy schools use standardised patients (SPs) in OSCEs, which increases the fidelity of the experience for students and enhances patient safety by allowing students to refine clinical skills prior to interacting with a real patient (Smithson *et al.,* 2015; Gillette *et al.,* 2017). Utilising simulations in pharmacy curricula also increases student confidence and prepares students for clinical practice (Korayem *et al.,* 2022).

OSCEs can be used for formative or summative assessment purposes (Sturpe, 2010; Deng, Fenn &Plake, 2019). Many schools use summative OSCEs to assess clinical and communication skills (Sturpe, 2010). Teaching OSCEs are sometimes used as a formative assessment method to enable students to learn from their mistakes in a low-stakes environment, identify strengths and weaknesses, and receive constructive feedback from evaluators (Deng, Fenn & Plake, 2019). Whether schools are implementing OSCEs for summative or formative purposes, it is essential to identify and understand differences in evaluator scoring and factors that could impact scoring. For summative OSCEs, this is relevant because the assessment should be valid and reliable, especially in high-stakes exams. Understanding these differences is also vital for teaching OSCEs to ensure that students are provided with accurate feedback to help them improve for future summative assessments and clinical practice (Chisnall *et al.,* 2015).

Additionally, while considered an effective method of assessment for clinical and communication skills, OSCEs are resource intensive, with such resources including rooms, equipment, faculty time and staff time, as well

as payment and training of SPs (Shirwaikar, 2015; Barrickman & Maynor, 2021). At West Virginia University School of Pharmacy, formative and summative OSCEs are primarily evaluated in real time by pharmacy faculty, requiring a significant time commitment from multiple faculty members for each OSCE done within the curriculum. The use of SPs as evaluators is a potential avenue for reducing the time needed for faculty to conduct OSCEs. According to one study, approximately 47% of pharmacy schools used SPs as evaluators (Sturpe, 2010). Other data have indicated that SP evaluators typically rated students higher than physician evaluators in a school of medicine, but little is known about comparisons of SP and faculty evaluator scoring in pharmacy OSCEs or factors that impact SP and faculty evaluation of student performance on OSCEs (McLaughlin *et al.,* 2006). Some available data suggest that with adequate training, SPs can accurately indicate what actions the student completed during an OSCE station (Ragan, Virtue & Chi, 2013). Data comparing SP and faculty evaluation of pharmacy student communication skills through the use of a standardised rubric appear to be lacking.

The objectives of the study were to (1) determine if SP assessment of student communication differs from faculty assessment and (2) identify factors that affect SP assessment, faculty assessment, and self-assessment of student communication skills.

Methods

Students

In spring 2022, all third-year PharmD students enrolled in a 4-year PharmD programme (N=69) participated in the described OSCEs as part of a required end-of-the-year capstone course. Table I presents student demographic information.

Each student completed four formative OSCEs, and each OSCE included two stations assessed by a faculty evaluator. For each OSCE, students were scheduled at different times on the same day. Students were not provided with information about the disease state topics that would be included in the assessment prior to the OSCE. During the OSCE stations, students were permitted to access electronic resources during the 8-minute encounter.

Faculty and standardised patients

Faculty and SP pairs were different for each OSCE station each week, with a total of 15 different faculty and 20 different SPs over the four weeks of OSCEs. Faculty involved in the OSCE were a combination of clinical and pharmaceutical sciences faculty involved in previous years of the curriculum. SPs were selected based on availability and trained on the cases by an SP educator and the faculty coordinators to minimise variations in SP performance.

Table I: Student demographics

Variables	Average (%)
Gender	
Female	41 (59.5%)
Male	28 (40.5%)
Ethnicity	
Asian	4 (5.7%)
Black	3 (4.4%)
Hispanic	1 (1.5%)
White	61 (88.4%)
	Average (SD)
Age (years)	23.6 (1.67)
Cumulative GPA	3.43 (0.43)
Term GPA	3.38 (0.57)

SD=standard deviation

OSCEs and students' assessment

Table II shows the topics for each OSCE station. For each case, students were evaluated on two different aspects: (1) clinical knowledge, using an analytical checklist that was unique to each OSCE case, and (2) communication skills, utilising a standard global assessment (GA) rubric. The GA rubric was consistent for all OSCE cases, independently of the scenario, and used a 5-point scale to assess the capability of a student to communicate in an organised manner, demonstrate genuine concern of patient's problems, have effective verbal and nonverbal communication, and display empathy.

Table II: OSCE station topics

OSCE	Station topic	Setting
1	Chronic Pain	Community Pharmacy
	Transitioning Anticoagulants	Ambulatory Clinic
2	Acute Kidney Injury	Hospital
	Type 2 Diabetes	Ambulatory Clinic
3	Immunisation Hesitancy	Community Pharmacy
	Hospital Acquired Pneumonia	Hospital
4	Epilepsy	Ambulatory Clinic
	Post-Traumatic Stress Disorder	Hospital

The analytical checklists for each station were only completed by a faculty evaluator that was observing the encounter remotely (not present in the room where the student and the SP were interacting), and the GA rubric for each station was completed by a faculty evaluator and an SP. The SP completed the GA rubric immediately following the encounter with each student. Faculty evaluators also had the opportunity to provide open-ended comments on student performance.

In addition to completing the GA rubric, each SP was also asked to rate their confidence in the student's recommendation on a scale of 1-5, with 1 being the lowest and 5 being the highest. This rating was completed immediately following each student encounter.

Finally, each student was asked to self-assess their performance using the GA rubric after each encounter and to rate their confidence on the recommendation that they provided to the SP on a scale of 1-5, with 1 being lowest and 5 being highest.

Data collection, data analysis, and statistical analysis

Data from each rubric were collected using the simulation centre's learning management system and downloaded in Prism and Excel. Skewness, kurtosis, and D'Agostino-Parson normality tests were calculated to identify Gaussian distribution among data sets (Prism 7.0e GraphPad Software®). The GA scores obtained from SP and faculty were compared. Additionally, the analytical checklist scores were compared with GA scores and confidence in the recommendations from SP and students. The Wilcoxon matched-pairs signed rank test was used for the analysis, and the ANOVA Friedman multiple analysis paired test was applied to assess the students' confidence across multiple days. Correlation analyses were performed using the Spearman R test (Prism 7.0e GraphPad Software®) to examine the relationship between students' confidence and the scores given by SPs or faculty. Correlations were categorised as weak (r = 0.10-0.29), moderate (r = 0.30-0.5), or strong (r > 0.5) based on the existing literature (Cohen, 1988).

Ethical considerations

As this study analysed OSCE assessment, which is a standard educational practice in a required pharmacy course, it was acknowledged as exempt by the Institutional Review Board.

Results

SP vs. faculty evaluator global assessment (GA) score

Students completed two faculty-graded OSCE stations each week for four weeks. Combined GA scores submitted by SPs in all OSCEs were statistically significantly higher than the GA scores submitted by faculty (4.4 ± .77, 95% CI 4.33 – 4.46 vs 4.1 ± .86, 95% CI 4.01 – 4.16, p<.001). This trend was maintained when analysing scores from individual OSCE days, as shown in Table III.

Table III: Summary of SP and Faculty Global Assessment Scores assigned to students

OSCE	GA Scores from SP Mean (SD) 95% CI	GA Scores from Faculty Mean (SD) 95% CI	*p*-value (Wilcoxon)
1	4.27 (0.83) 4.13 – 4.41	4.1 (0.87) 3.95 – 4.25	0.008
2	4.3 (0.79) 4.16 – 4.43	4.11 (1.01) 3.93 – 4.28	0.05
3	4.49 (0.6) 4.39 – 4.59	4.2 (0.78) 4.07 – 4.33	<0.001
4	4.52 (0.82) 4.38 – 4.66	3.92 (0.765) 3.79 – 4.06	<0.001

GA=global assessment; SD=standard deviation; CI=confidence interval

Correlation of scoring

Correlations were analysed to determine associations between SP global assessment scores, SP confidence in student recommendations, student GA scores, student self-confidence in recommendations, faculty evaluator GA scores, and student performance on the analytical checklist component of each OSCE station. Results are shown in Table IV.

Table IV: Correlation of OSCE evaluation components from students, faculty, and standardised patients (Spearman R test)

OSCE		GA SP	SP confidence	GA student	Student confidence	GA faculty
1	**SP confidence**	0.78***				
	GA student	0.05	0.12			
	Student confidence	0.07	0.19	0.49***		
	GA faculty	0.49***	0.51***	0	0.17*	
	Analytical checklist	0.43***	0.41***	0.07	0.16	0.34***
2	**SP confidence**	0.79***				
	GA student	0.25**	0.26**			
	Student confidence	0.41***	0.42***	0.69***		
	GA faculty	0.41***	0.57***	0.28***	0.36***	
	Analytical checklist	0.43***	0.48***	0.26***	0.34***	0.63***
3	**SP confidence**	0.70***				
	GA student	0	0.13			
	Student confidence	0	0.07	0.62***		
	GA faculty	0.14	0.19*	0.05	0.02	
	Analytical checklist	0.07	0.12	0.17*	0.31***	0.12
4	**SP confidence**	0.75***				
	GA student	0.05	0.14			
	Student confidence	0.05	0.12	0.61***		
	GA faculty	0	0.07	0.20*	0.18*	
	Analytical checklist	0	0.12	0.12	0.13	0.51***

OSCE=objective structured clinical examination; SP=standardized patient; GA=global assessment
*p-values are shown for statistically significant correlations: *p<0.05; **p<0.01; ***p<0.001*

Faculty evaluator GA score vs analytical checklist score

For most OSCEs, there was a moderate-high correlation between the faculty GA score and student performance on the analytical checklist, exception for OSCE 3. No strong correlations were identified between faculty GA score and student GA score or between faculty GA score and student confidence in their recommendations.

Confidence scores

Combined confidence scores submitted by SPs in all OSCEs were statistically significantly higher than the self-confidence scores submitted by students (4.4 ± .84, 95% CI 4.29 – 4.43 vs 3.1 ± 1.07, 95% CI 3.03 – 3.21, p<.001). This trend was maintained when analysing scores from individual OSCE days, as shown in Table V.

SP confidence was not correlated with students' self-confidence, except for one case on OSCE 2, which focused on Type 2 diabetes, as shown in Table V. Furthermore, SP GA scores and confidence in student recommendations were strongly correlated, independent of the case topic or setting. In most scenarios, the SP GA score was not correlated with the student's self-assessed GA score or student self-confidence, with one exception during OSCE day 2, where there was a direct correlation between both scores.

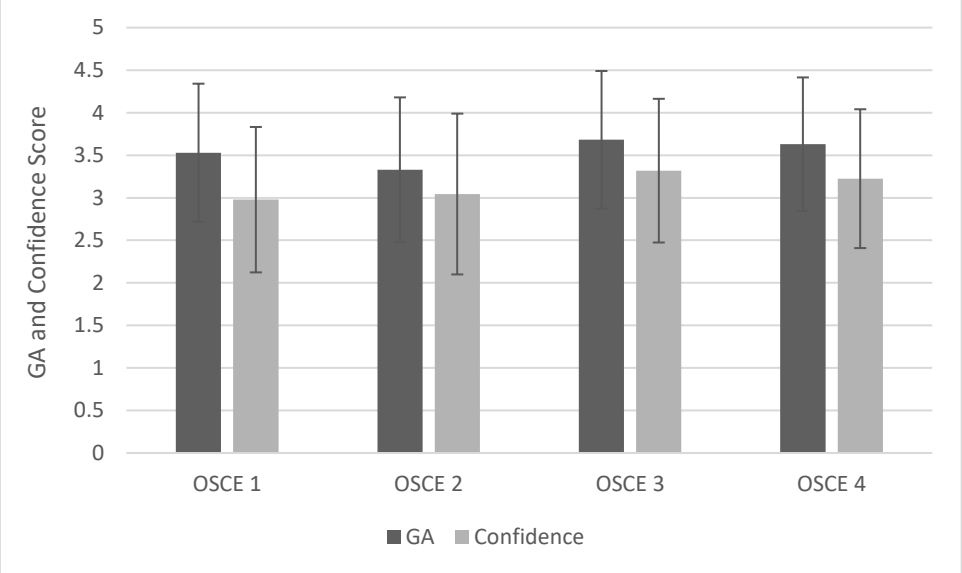
Table V: Summary of confidence scores assigned by standardised patients and self-confidence scores provided by students

OSCE	SP confidence Mean (SD) 95% CI	Student confidence Mean (SD) 95% CI	*p*-value (Wilcoxon)
1	4.08 (0.99) 3.91 – 4.25	2.96 (1.09) 2.78 – 3.15	p<0.001
2	4.33 (0.81) 4.19 – 4.47	3.01 (1.11) 2.89 – 3.2	p<0.001
3	4.53 (0.63) 4.42 - 4.64	3.32 (0.94) 3.52 – 3.84	p=0.001
4	4.50 (0.83) 4.36 – 4.65	3.21 (0.97) 3.04 – 3.38	p<0.001

OSCE=objective structured clinical examination; SD=standard deviation; CI=confidence interval

Student confidence

Since these formative OSCEs were offered over four weeks, an analysis comparing student self-confidence across days was performed to determine if confidence increased as students became more familiar with OSCEs. Figure 1 shows the evolution of student self-confidence and self-assessed GA scores throughout the formative OSCE experiences. Friedman multiple analysis paired test revealed statistical significance when comparing average student self-assessed GA for OSCE 2 vs OSCE 3 (p=0.004). The comparison of student confidence between OSCE 1 and OSCE 3 was also statistically significant (p=0.004).



OSCE=Objective structured clinical examination; GA=Global assessment

Figure 1: Student self-assessed global assessment scores and confidence

Discussion

The primary objective of this study was to determine if there was a difference in OSCE communication scores between SPs and faculty evaluators using a standard GA rubric. In our cohort, there was a statistically significant difference in the assessment of student communication, with SP evaluators assigning higher scores on the standardised rubric compared with the score given by faculty. This finding is consistent with previous data on medical students (McLaughlin *et al.,* 2006). As hypothesised by Schwartzman and colleagues, differences in communication skills between SPs and faculty can occur because the assessment of communication skills is inherently subjective, and faculty may have more experience evaluating with rubrics compared to SPs (Schwartzman *et al.,* 2011).

Our analysis also found a strong correlation between SP confidence in student recommendations and SP GA scores. Interestingly, our institution does not provide SPs with the analytical checklist for OSCE cases, and most SPs do not have any formal healthcare training. Therefore, many do not know the correct "clinical" answer to the case. Although not addressed in the published literature, the authors speculate that while confidence in a specific recommendation is not explicitly linked to communication, it stands to reason that students who are more skilled communicators may instil confidence in the SP regarding their recommendations, irrespective of whether or not the student recommendation was clinically correct.

As indicated in previous studies, the need for OSCE evaluator training is essential and can be a disadvantage in implementing OSCEs (Salinitri *et al.,* 2011; Shirwaikar, 2015). Further analysis indicated that faculty scores on the overall assessment sometimes, but not always, correlated with students' performance on the analytical checklist, despite evaluators' training to assess communication skills separately from clinical knowledge/performance. This finding suggests that OSCE evaluators may benefit from additional training to ensure that evaluations of clinical knowledge and communication skills as distinct performance measures are valid.

Interestingly, student confidence only showed mild correlations with faculty global assessment scores, and only one of four OSCEs showed a moderate correlation between the SP global assessment score and student confidence. A common concern among faculty at our institution is that poor performance on the first OSCE station could impact student confidence and thus hinder performance on subsequent OSCE stations; however, our results indicate that this is not necessarily the case, as student confidence was also not correlated to performance on the analytical checklist. This finding contrasts with published data on physical therapy students, showing that student confidence was associated with adequate performance on OSCEs (Ferreira *et al.,* 2020).

Student confidence consistently improved with each OSCE, and SP confidence in student recommendations also increased from the first OSCE to the third OSCE, demonstrating the adage "practice makes perfect" and supporting the use of OSCEs throughout pharmacy curricula to improve student confidence before advanced pharmacy practice experiences (APPEs) and clinical practice.

Limitations

Strengths of the study include the analysis of eight OSCE stations over four different weeks and the use of more than 20 different SPs and 15 different faculty evaluators.

Limitations of the study include implementation at a single institution with one cohort of students and potential evaluator bias, which the authors attempted to minimise by randomly assigning students, evaluators, and SPs for each OSCE case. Future studies could assess additional factors that may influence subjective scoring in OSCEs or student confidence on OSCE performance.

Conclusion

Evaluating communication skills involves subjective perception, and it is crucial to determine variations in OSCE scoring among evaluator types, especially as more pharmacy schools incorporate OSCEs as the gold standard for assessing student clinical and communication skills. Possible disparities in evaluating communication skills in OSCEs and factors that affect scoring should be accounted for when selecting evaluators for high-stakes OSCEs. It is also essential to ensure that students are receiving reliable feedback from evaluators in formative OSCEs.

Conflict of interest

The authors declare no conflict of interest to disclose.

Source of funding

The authors did not receive any funding.

Acknowledgement

The authors would like to acknowledge the David and Jo Ann Shaw Center for Simulation Training and Education for Patient Safety (STEPS Center) at West Virginia University for their assistance with implementing OSCEs.

References

Barrickman, A. L., & Maynor, L. (2021). PROGRAMME DESCRIPTION: Implementation of a capstone course with Formative OSCEs to evaluate preparedness for Advanced Pharmacy Practice Experiences (APPE). *Pharmacy Education,* **21**, 145–158. https://doi.org/10.46542/pe.2021.211.145158

Chisnall, B., Vince, T., Hall, S., & Tribe, R. (2015). Evaluation of outcomes of a formative objective structured clinical examination for second-year UK medical students. *International journal of medical education*, **6**, 76–83. https://doi.org/10.5116/ijme.5572.a534

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Deng, B., Fenn, N. E., 3rd, & Plake, K. S. (2019). Impact of a teaching objective structured clinical examination (TOSCE) on student confidence in a pharmacy skills laboratory. *Currents in pharmacy teaching & learning*, **11**(2), 145–154. https://doi.org/10.1016/j.cptl.2018.11.009

Ferreira, É. M. R., Pinto, R. Z., Arantes, P. M. M., Vieira, É. L. M., Teixeira, A. L., Ferreira, F. R., & Vaz, D. V. (2020). Stress, anxiety, self-efficacy, and the meanings that physical therapy students attribute to their experience with an objective structured clinical examination. *BMC medical education*, **20**(1), 296. https://doi.org/10.1186/s12909-020-02202-5

Gillette, C., Stanton, R. B., Rockich-Winston, N., Rudolph, M., & Anderson, H. G., Jr (2017). Cost-Effectiveness of Using Standardized Patients to Assess Student-Pharmacist Communication Skills. *American journal of pharmaceutical education*, **81**(10), 6120. https://doi.org/10.5688/ajpe6120

Korayem, G. B., Alshaya, O. A., Kurdi, S. M., Alnajjar, L. I., Badr, A. F., Alfahed, A., & Cluntun, A. (2022). Simulation-Based Education Implementation in Pharmacy Curriculum: A Review of the Current Status. *Advances in medical*

education and practice, **13**, 649–660. https://doi.org/10.2147/AMEP.S366724

Kristina, S. A., & Wijoyo, Y. (2019). Assessment of Pharmacy Students? Clinical Skills using Objective Structured Clinical Examination (OSCE): A Literature Review. *Systematic Reviews in Pharmacy,* **10**(1), 55–60. https://doi.org/10.5530/srp.2019.1.9

McLaughlin, K., Gregor, L., Jones, A., & Coderre, S. (2006). Can standardized patients replace physicians as OSCE examiners?. *BMC medical education*, **6**, 12. https://doi.org/10.1186/1472-6920-6-12

Ragan, R. E., Virtue, D. W., & Chi, S. J. (2013). An assessment program using standardized clients to determine student readiness for clinical practice. *American journal of pharmaceutical education*, **77**(1), 14. https://doi.org/10.5688/ajpe77114

Salinitri, F. D., O'Connell, M. B., Garwood, C. L., Lehr, V. T., & Abdallah, K. (2012). An objective structured clinical examination to assess problem-based learning. *American journal of pharmaceutical education*, **76**(3), 44. https://doi.org/10.5688/ajpe76344

Schwartzman, E., Hsu, D. I., Law, A. V., & Chung, E. P. (2011). Assessment of patient communication skills during OSCE: examining effectiveness of a training program in minimizing inter-grader variability. *Patient education and counseling*, **83**(3), 472–477. https://doi.org/10.1016/j.pec.2011.04.001

Shirwaikar A. (2015). Objective structured clinical examination (OSCE) in pharmacy education - a trend. *Pharmacy practice*, **13**(4), 627. https://doi.org/10.18549/PharmPract.2015.04.627

Smithson, J., Bellingan, M., Glass, B., & Mills, J. (2015). Standardized patients in pharmacy education: An integrative literature review. *Currents in Pharmacy Teaching and Learning,* **7**(6), 851–863. https://doi.org/10.1016/j.cptl.2015.08.002

Sturpe D. A. (2010). Objective structured clinical examinations in doctor of pharmacy programs in the United States. *American journal of pharmaceutical education*, **74**(8), 148. https://doi.org/10.5688/aj7408148

Urteaga, E. M., Attridge, R. L., Tovar, J. M., & Witte, A. P. (2015). Evaluation of Clinical and Communication Skills of Pharmacy Students and Pharmacists with an Objective Structured Clinical Examination. *American journal of pharmaceutical education*, **79**(8), 122. https://doi.org/10.5688/ajpe798122

Weaver, S. B., Daftary, M., Wingate, L., & Turner, M. (2022). Evaluation of faculty inter-variability OSCE grade scoring on overall student performance in a laboratory course. Pharmacy Education, 22(1), 48–53. https://doi.org/10.46542/pe.2022.221.4853