

RESEARCH ARTICLE

Association between Objective Structured Clinical Examination (OSCE) station order and clinical competence assessment in pharmacy students

Lena Maynor , Ashleigh Barrickman , Marina Gálvez-Peralta 

West Virginia University School of Pharmacy, West Virginia, United States

Keywords

Clinical skills assessment
OSCE
Pharmacy student
Student performance

Correspondence

Lena Maynor
West Virginia University School of
Pharmacy
West Virginia
United States
lmaynor@hsc.wvu.edu

Abstract

Background: Objective structured clinical examinations (OSCEs) are used to assess clinical skills in pharmacy education. Some concern exists regarding station order and OSCE performance, but data are lacking in pharmacy students. The purpose of this study was to determine whether OSCE station order or first-station student confidence are associated with overall OSCE performance in third-year pharmacy (P3) students. **Methods:** This retrospective observational cohort study included P3 students completing four three-station formative OSCEs and a two-day summative OSCE with four stations per day. Stations were ranked by difficulty using pre-determined cut scores. For formative OSCEs, students also reported immediate post-station confidence. Overall OSCE performance was compared between students completing the most difficult station first versus all others and between low and neutral-high first station student confidence groups using the Mann-Whitney test. **Results:** Completing the most difficult station first was not associated with overall formative or summative OSCE performance. However, students reporting low confidence on their first station had lower overall formative OSCE performance. Student confidence was higher on stations designated as most difficult. **Conclusion:** OSCE station order does not appear to impact overall performance, but low student confidence on the initial OSCE station may be associated with overall lower performance.

Introduction

The objective structured clinical examination (OSCE) has been used to assess clinical and communication skills extensively in health professions education, including pharmacy, and is considered a standard assessment tool in many curricula. The use of OSCEs in pharmacy has expanded over the last decade, with more than 50% of accredited programmes in the United States using them in their curricula (Barnes *et al.*, 2024). The 2025 Accreditation Council for Pharmacy Education (ACPE) standards for programmes in the United States emphasise the importance of pharmacy graduates having adequate problem-solving and critical thinking skills, and OSCEs have been shown to be more effective than traditional assessment methods at assessing these essential skills (Austin *et al.*, 2003; Barnes *et al.*, 2024).

Due to the widespread use of OSCEs in pharmacy curricula, the summative nature of OSCEs in many programmes, and the need to assess problem-solving and critical thinking skills, it is important to identify factors that can influence student performance in an OSCE.

Potential factors that can impact student performance in an OSCE include variations in case portrayal by standardised patients, inter-evaluator variability, and levels of student experience (Ferreira *et al.*, 2020; Weaver *et al.*, 2022). As with other types of examinations, OSCEs may produce high levels of stress and anxiety in health professions students (Zhang & Rabatsky, 2015; Ferreira *et al.*, 2020; Martin & Naziruddin, 2020; Miranda & de la Barra Ortiz, 2023; Massalou *et al.*, 2025). However, neither stress nor anxiety have been consistently associated with OSCE

performance (Zhang & Rabatsky, 2015; Ferreira *et al.*, 2020; Martin & Naziruddin, 2020; Miranda & de la Barra Ortiz, 2023; Massalou *et al.*, 2025). Similarly, student self-efficacy or perception of performance may be inconsistent with actual performance (Mavis, 2001; Morgan & Cleave-Hogg, 2002; Chan *et al.*, 2015; Lim *et al.*, 2024).

The logistics of an OSCE typically include multiple timed stations that require students to interact with a standardised patient, with students rotating through stations in varied orders. Some concern in the literature exists regarding the effects of early station performance on subsequently completed stations within an individual OSCE (Monteiro *et al.*, 2016; Bellido-Esteban *et al.*, 2021). One such concern is related to working memory, which is utilised in skill-focused activities. Working memory is defined as "the small amount of information that can be held in mind and used in the execution of cognitive tasks" (Cowan, 2014). Anxiety or high-pressure has been shown to occupy working memory capacity, which can limit the total capacity that can be devoted to skill execution. The occupation of working memory capacity has been associated with "choking under pressure," or uncharacteristic underperformance in high stakes situations (Beilock & Carr, 2005). One study of OSCE performance in psychology students found that students completing high complexity stations first in the OSCE rotation had higher autonomic stress responses than students completing low complexity stations first, based on heart rate monitoring and R-R heartbeat intervals (Bellido-Esteban *et al.*, 2021). Another study reviewing station performance to determine difficulty found no difference in medical student performance based upon station order (Monteiro *et al.*, 2016). Data on the effect of OSCE station order on student performance appear to be lacking in pharmacy students.

The objective of this study was to determine the impact of OSCE station order on overall OSCE performance in third-year pharmacy students.

Methods

This project was approved by the Institutional Review Board of the academic institution where the study was conducted. Third-year pharmacy (P3) students completed four three-station teaching (formative) OSCEs and a two-day summative OSCE (four stations per day) during a capstone course in the spring 2022 semester. Scores from all students who completed the OSCEs on their scheduled dates were included in the analysis. OSCE stations were ranked by difficulty using

cut scores, which were predetermined using a modified Angoff method (Jalili *et al.*, 2011). During this iterative process, faculty are asked to rate the percentage of "minimally competent" students, i.e. those students who possess the lowest acceptable level of knowledge and skills, who would complete each item on the station checklist (Boursicot & Roberts, 2006). Faculty rankings are averaged for each line item, and the sum of the averages becomes the station cut score. Overall teaching OSCE scores were calculated by adding the scores of each station. All station scores were on a 100-point scale, and the total number of overall OSCE points for all stations combined was 300 for the teaching OSCEs. Similarly, station scores for each day of the summative OSCE were added to determine that day's score, with 400 total available points for each day.

For teaching OSCEs, students rated their confidence in their recommendations on a scale from 1 (lowest) to 5 (highest) immediately after completing each station. Individual performance data and confidence were paired for each student. If any station within the relevant OSCE lacked a confidence rating, the student performance data were excluded from the confidence analysis. Due to increased security measures and scheduling constraints within the summative OSCEs, students were not asked to rate their confidence on stations during the summative OSCE. Overall teaching OSCE performance was compared between students who completed the most difficult station first and the students who completed any other station first. Similarly, performance on formative OSCEs was compared between students who rated their confidence on the first station as 1-2 (low) versus 3-5 (neutral-high). Additionally, student confidence for the most difficult station was compared to student confidence for all other stations.

For the summative OSCE, each day of the summative OSCE was treated as its own exam for purposes of data analysis related to station order, and comparison was made of overall performance for students who completed the most difficult OSCE station first compared to performance of all other students. An additional comparison of overall summative OSCE performance (both days combined) was completed based on average reported confidence on prior teaching OSCEs. All comparisons were made using the Mann-Whitney test, GraphPad Prism version 10.4.1.

Ethics approval and informed consent

This study was approved by the West Virginia University Institutional Review Board (IRB Protocol No. 1304035754; Project No. 22022402). The requirement for informed consent was waived.

Results

Sixty-nine students completed the capstone course in spring 2022, with 274 instances of teaching OSCEs completed on their originally scheduled date. Student demographic information can be found in Table I. All students completed both days of the summative OSCE as scheduled, with 138 overall scores included in the analysis. A summary of teaching OSCE station topics, station cut scores, average station score, and average station confidence can be found in Table II, and a summary of summative OSCE station topics, cut scores, and average station scores can be found in Table III.

Table I: Demographic Characteristics of Study Participants (N=69)

Characteristic	Value
Sex, n (%)	
Female	42 (60.9)
Male	27 (39.1)
Age, average years (range)	24 (23-31)
Cumulative GPA*, average (SD [†])	3.48 (0.45)

*GPA = grade point average

[†]SD = standard deviation

Table II: Summary of teaching OSCE stations*

OSCE	Station topic	Cut score	Average student score	Average student confidence
1	Chronic pain	60	69	3.19
	Tobacco cessation	48	80	3.96
	Anticoagulation	59	75	2.73
2	Acute kidney injury	57	40	2.65
	Type 2 diabetes mellitus	52	73	3.39
	Hypertension	59	78	3.38
3	Vaccines	67	80	3.38
	Osteoarthritis	51	71	2.97
	Hospital acquired pneumonia	55	72	3.67
4	Drug induced constipation	54	75	3.03
	Epilepsy	49	58	3.63
	Post-traumatic stress disorder	59	67	3.39

*All station cut scores scaled to 100 points. Student confidence was scored on a scale of 1 (lowest) to 5 (highest)

Table III: Summary of summative OSCE stations*

OSCE day	Station topic	Cut score	Average student score
1	Asthma	59	84
	Resistant hypertension	49	69
	Chronic kidney disease	52	71
	Type 2 diabetes mellitus	60	74
2	Community acquired pneumonia	56	58
	Heartburn	60	82
	Migraine	48	68
	Depression	59	84

*All station cut scores scaled to 100 points

For teaching OSCEs, scores were similar between students who completed the most difficult station first compared to students who started with any other station (216.1 ± 29.54 vs. 211.3 ± 32.66 , $p=0.28$), as shown in Figure 1. Similarly, performance was similar on the summative OSCE for students who completed

the most difficult station first compared to other students (292.7 ± 50.13 vs. 295.4 ± 33.09 , $p=0.87$), as shown in Figure 2. In contrast, students who reported low confidence on their first teaching OSCE station performed worse overall on that same teaching OSCE compared to students reporting neutral to high

confidence on their first OSCE station (202.5 ± 33.73 vs. 215.6 ± 30.59 , $p < 0.001$), as shown in Figure 3. There was no difference in overall summative OSCE performance (both days combined) when comparing students with an average confidence rating of low (1-2)

vs neutral-high (3-5) on prior teaching OSCEs (580.8 ± 57.63 vs 597.6 ± 48.92 , $p = 0.33$).

Student-rated confidence in recommendations was higher on the most difficult stations compared to other stations (3.48 ± 1.03 vs. 3.18 ± 1.07 , $p < 0.0001$).

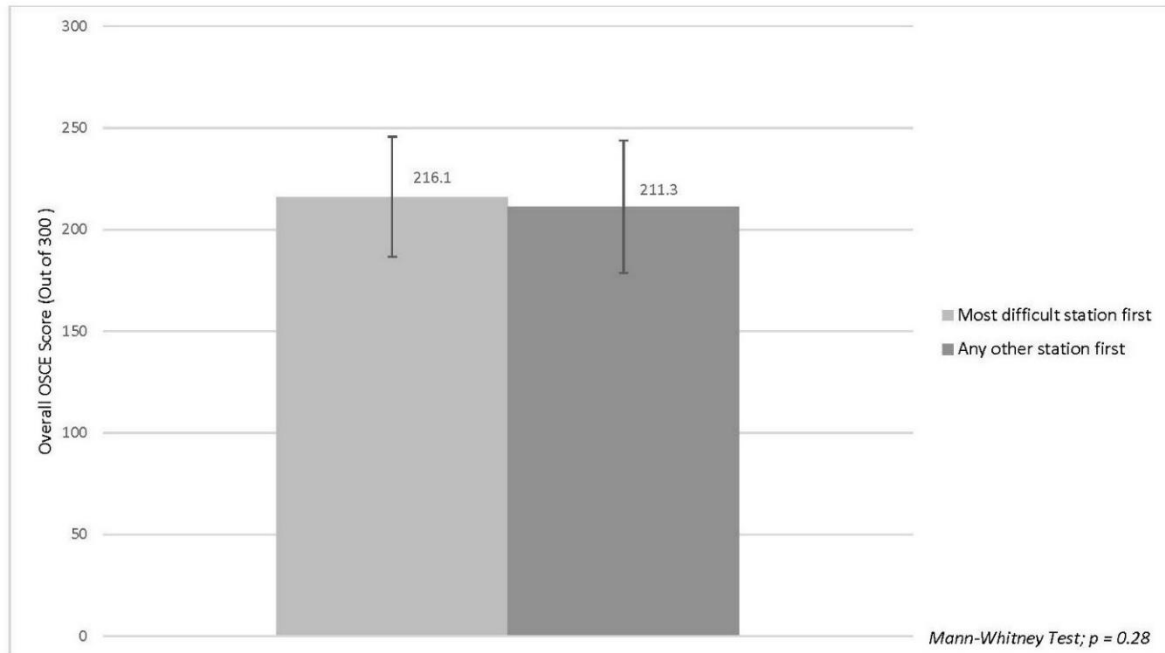


Figure 1: Average overall teaching OSCE performance by difficulty of first station.

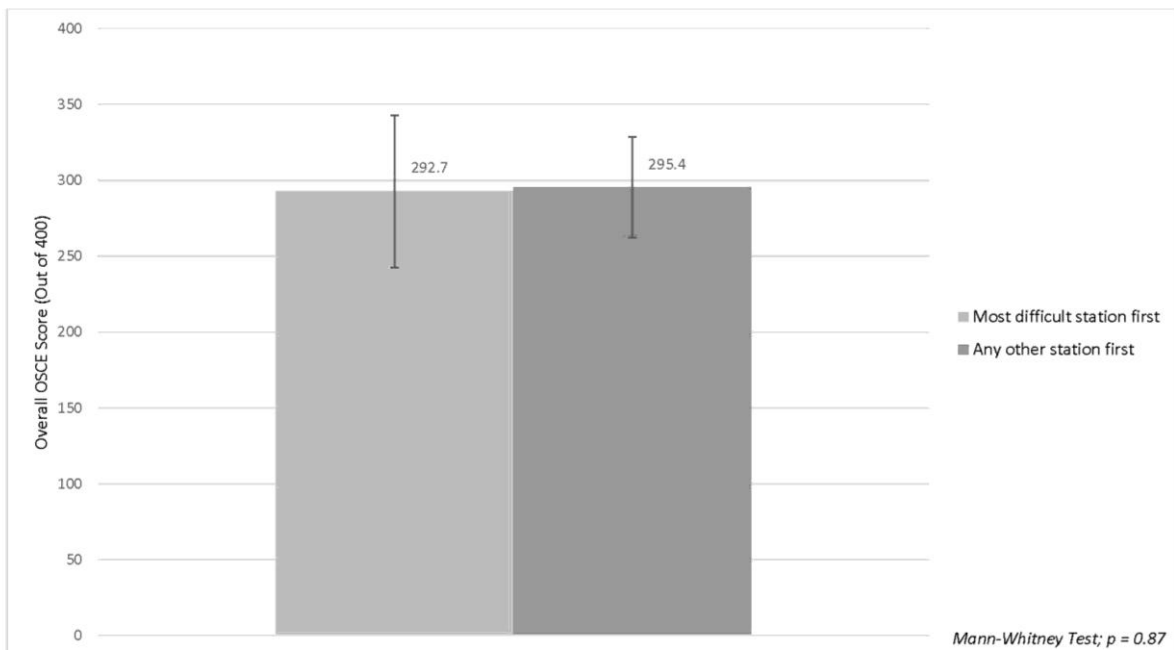


Figure 2: Average overall summative OSCE performance by difficulty of first station.

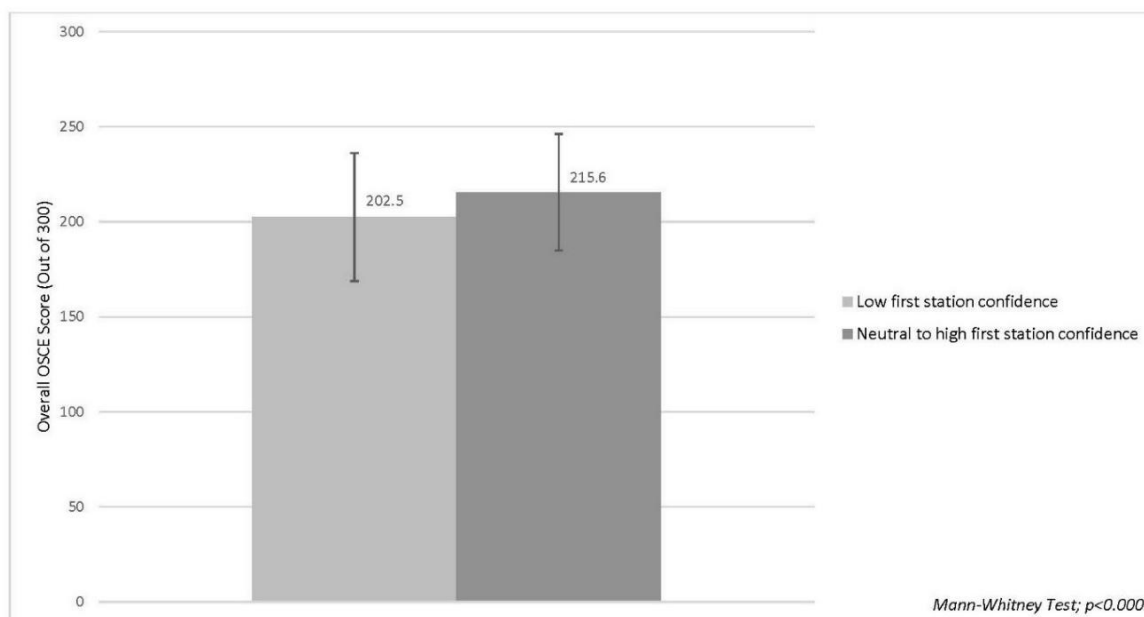


Figure 3: Average overall performance on teaching OSCE by student confidence on the first station.

Discussion

Understanding factors that can affect student performance on an OSCE is imperative in pharmacy education, especially as OSCEs become more widely used across the academy. Completing the most difficult station first during an OSCE was not associated with overall OSCE performance within this cohort. This finding is consistent with those found in medical students (Monteiro *et al.*, 2016). The structure of the OSCE requires students to move through stations in varied orders to allow the OSCE to be delivered in an efficient manner. These results provide some reassurance that this structure does not create a disadvantage when a student is assigned to complete the most difficult station first.

In contrast, students who had low confidence in their recommendation on the first station during a teaching OSCE had a lower overall score on that same teaching OSCE compared to students who had neutral or high confidence in their first station, independent of the level of difficulty of the station. This is different from most findings that indicate that while stress and anxiety may be elevated during an OSCE, neither stress nor anxiety has been shown to affect OSCE performance (Zhang & Rabatsky, 2015; Ferreira *et al.*, 2020; Martin & Naziruddin, 2020; Miranda & de la Barra Ortiz, 2023; Massalou *et al.*, 2025). While stress and anxiety were not measured as part of this study, it is reasonable to infer that anxiety and stress are connected to low confidence in performance (Chakraborty, 2023). It is also

possible that students who were not confident in their recommendation are also students who would have performed more poorly regardless of station order or who are facing more challenges with working memory, where hierarchy of steps, processes, and interconnection are relevant (Cowan, 2014). Interestingly, average reported confidence by station was sometimes, but not always, aligned with average student performance in our cohort. For example, for teaching OSCE 3, the average score for the osteoarthritis station was 71 with an average reported confidence of 2.97, while the average score on the hospital acquired pneumonia station was 72 with an average reported confidence of 3.67. Another study found that pharmacy student self-reflections on OSCE performance generally did not align with actual scores. In that study, most of the discrepancies came from high performing students judging themselves harshly, while poorer performing students' self-reflections were more generally in-line with their overall scores (Lim *et al.*, 2024). Notably, reported student confidence on the most difficult station, as defined by cut score, was higher than for other stations. Cut scores are determined by a modified Angoff method (Jalili *et al.*, 2011) to a minimal competence standard and are expected to be lower than average student performance. In general, a lower cut score reflects an expectation that a "minimally competent student" would complete fewer checklist items, suggesting greater overall station difficulty. However, cut scores are determined for each station independently and might not reflect the relative difficulty of individual stations within the same OSCE.

Average student performance was usually higher than the cut score but not always aligned with our assigned difficulty of each station, as shown in Tables II and III. Some stations with the lowest cut scores, such as tobacco cessation and type 2 diabetes, are topics for which students receive multiple opportunities for reinforcement prior to the OSCE. It is also possible for these stations in particular, students felt more confident in their recommendation because of their level of exposure to the topic itself, and repetition facilitated the transition from working memory capacity to long-term memory (Cowan, 2014).

Limitations

This study has several limitations. It was conducted at a single institution during one academic semester and was retrospective, using existing performance and student confidence data, which may limit generalisability. Students were not randomised into station order, and other potential confounders of performance were not controlled. Additionally, stress and anxiety were not directly measured. Due to more complex logistics, student confidence was not assessed during the summative OSCE. It is possible that student confidence may be impacted by the high-stakes nature of the summative OSCE. This is an area of potential future study.

Conclusion

While station order itself may not impact OSCE performance, students with low confidence on their initial OSCE station may benefit from monitoring and intervention to build confidence in recommendations.

Conflict of interest

The authors declare no conflict of interest.

Source of funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Austin, Z., O'Byrne, C., Pugsley, J., & Quero Munoz, L. (2003). Development and validation processes for an objective structured clinical examination (OSCE) for entry-to-practice certification in pharmacy: the Canadian experience. *American Journal of Pharmaceutical Education*, *67*(3), 76. [https://doi.org/10.1016/S0002-9459\(24\)00664-8](https://doi.org/10.1016/S0002-9459(24)00664-8)
- Barnes, K.N., Hardinger, K.L., Graham M.R. (2024). Review of objective structured clinical examination practices within pharmacy programs throughout the United States. *American Journal of Pharmaceutical Education*, *88*(4), 100686. <https://doi.org/10.1016/j.ajpe.2024.100686>.
- Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: working memory and “choking under pressure” in math. *Psychological Science*, *16*(2), 101–105. <https://doi.org/10.1111/j.0956-7976.2005.00789.x>
- Bellido-Esteban, A., Beltrán-Velasco, A. I., Ruisoto-Palomera, P., Nikolaidis, P. T., Knechtle, B., & Clemente-Suárez, V. J. (2021). The effect of psychology objective structured clinical examination scenarios presentation order on students autonomic stress response. *Frontiers in Psychology*, *12*, 622102. <https://doi.org/10.3389/fpsyg.2021.622102>
- Boursicot K., & Roberts L. (2006). Setting standards in a professional higher education course: defining the concept of the minimally competent student in performance-based assessment at the level of graduation from medical school. *Higher Education Quarterly*, *60*(1), 74–90. <https://doi.org/10.1111/j.1468-2273.2006.00308.x>
- Chakraborty, A. (2023). Exploring the root causes of examination anxiety: effective solutions and recommendations. *International Journal of Science and Research*, *12*(2), 1096–1102. <https://doi.org/10.21275/sr23220002911>
- Chan, M., Bax, N., Woodley, C., Jennings, M., Nicolson, R., & Chan, P. (2015). The first OSCE; Does students' experience of performing in public affect their results? *BMC Medical Education*, *15*, 59. <https://doi.org/10.1186/s12909-015-0343-0>
- Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*(2), 197–223. <https://doi.org/10.1007/s10648-013-9246-y>
- Ferreira, É. D. M. R., Pinto, R. Z., Arantes, P. M. M. H., 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*, 296. <https://doi.org/10.1186/s12909-020-02202-5>
- Jalili, M., Hejri, S.M., & Norcini, J.J. (2011). Comparison of two methods of standard setting: the performance of the three-level Angoff method. *Medical Education*, *45*(12), 1199–1208. <https://doi.org/10.1111/j.1365-2923.2011.04073.x>
- Lim, A. S., Krishnan, S., Tan, G., Stewart, D., & Al-Diery, T. (2024). Do students' self-reflections of performance align with their graded performance in objective structured clinical exams? *Currents in Pharmacy Teaching and*

Learning, **16**(8), 102097.

<https://doi.org/10.1016/j.cptl.2024.04.013>

Martin, R. D., & Naziruddin, Z. (2020). Systematic review of student anxiety and performance during objective structured clinical examinations. *Currents in Pharmacy Teaching and Learning*, **12**(12), 1491–1497.

<https://doi.org/10.1016/j.cptl.2020.07.007>

Massalou, D., Doyen, J., Almairac, F., Rohrllich, P.-S., Fournier, J.-P., Vandersteen, C., & Sadaghianloo, N. (2025). Can relaxation exercises improve students' OSCE grades: a prospective study. *Canadian Medical Education Journal*, **16**(2), 43–47. <https://doi.org/10.36834/cmej.77739>

Mavis, B. (2001). Self-efficacy and OSCE performance among second year medical students. *Advances in Health Sciences Education*, **6**(2), 93–102.

<https://doi.org/10.1023/a:1011404132508>

Miranda, L. G., & de la Barra Ortiz, H. A. (2023). Anxiety and performance in a structured objective clinical examination of undergraduate physical therapy students. *Journal of Advanced Pharmacy Education and Research*, **13**(4), 83–92.

<https://doi.org/10.51847/ZQFM8JUZVG>

Monteiro, S. D., Walsh, A., & Grierson, L. E. M. (2016). OSCE circuit performance effects: Does circuit order influence scores? *Medical Teacher*, **38**(1), 98–100.

<https://doi.org/10.3109/0142159X.2015.1075647>

Morgan, P. J., & Cleave-Hogg, D. (2002). Comparison between medical students' experience, confidence and competence. *Medical Education*, **36**(6), 534–539.

<https://doi.org/10.1046/j.1365-2923.2002.01228.x>

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>

Zhang, N., & Rabatsky, A. (2015). Effects of test stress during an objective structured clinical examination. *Journal of Chiropractic Education*, **29**(2), 139–144.

<https://doi.org/10.7899/jce-14-17>