Pharmacy Education (2023) 23(1) 533 - 538
https://doi.org/10.46542/pe.2023.231.533538

PROGRAMME DESCRIPTION

An objective structured practical examination for laboratory skills in a pharmacy technician programme

Peter A C McPherson[1,2], Ben M Johnston[2]

[1] School of Pharmacy & Pharmaceutical Science, Ulster University, Coleraine, United Kingdom
[2] Belfast Metropolitan College, Titanic Quarter Campus, Belfast, United Kingdom

Keywords
Assessment
Borderline regression model
Laboratory skill
Observation

Correspondence
Peter A C McPherson
School of Pharmacy & Pharmaceutical Science
Ulster University
Coleraine
United Kingdom
p.mcpherson@ulster.ac.uk

Abstract

Background: Pharmacy technician education is typically at the pre-degree level and comprises instruction in scientific and clinical disciplines. The assessment of practical laboratory skills often utilises attainment-referenced methods, which are not always appropriate for vocationally-focused programmes. **Methods:** An objective structured practical examination (OSPE) was introduced to assess student competency in three key areas (accurate weighing, calibrating a pH meter and performing a dilution). Students were assessed using weighted criterion-based assessment criteria and an overall global performance rating, which allowed cut scores to be determined using a borderline regression method. Student opinions were collected using online questionnaires on a five-point Likert scale. **Results:** The move to OSPEs did not significantly alter the distribution of student results from previous years (mean ± SD, OSPE *vs* legacy: 77 ± 19% *vs* 73 ± 21%), suggesting that academic integrity was maintained. There was a high level of consistency in Likert score responses (Cronbach's α = 0.823), with students clearly favouring the OSPE approach. **Conclusion:** The move to an OSPE-based assessment was successful and provided a basis for the development of similar assessment strategies in the pharmacy technician programme.

Introduction

Within any pharmacy programme, chemistry stands as a fundamental component of the early curriculum, bringing forth an array of laboratory activities that develop psychomotor skills and lay the foundations for more advanced work in pharmaceutics, extemporaneous dispensing, and aseptic manipulations (Sosabowski & Gard, 2008). The assessment of laboratory work within the chemistry curriculum at this level often involves a practical examination encompassing a blend of activities and recall-based interpretation of data (Gott & Duggan, 2002). This assessment approach may not necessarily reflect authentic evaluation, as it can be possible to achieve a passing grade without demonstrating practical competency (Gericke *et al.*, 2022). For example, in a question where students were asked to describe and carry out a dilution, they could attain a passing grade by providing a sound theoretical explanation of the procedure and performing accurate calculations; hence, the practical aspect could be ignored. A common alternative is to assess learners through written submissions such as laboratory reports; however, for large cohorts, it is impractical to mark written submissions consistently, even with rubrics and moderation between assessors. The impact of so-called "*essay mills*", or more recently artificial intelligence chatbots such as ChatGPT, make it increasingly difficult to discern the authenticity of a student's written submission even with the assistance of anti-plagiarism software (Lee, 2023).

A further alternative is to use criterion-based assignments, where students must satisfy specific assessment criteria rather than achieve a passing mark based on relative standards (Newton, 2011). This method has clear advantages for professions where competency must be demonstrated for reasons of public safety (Pereira *et al.,*

ISSN 1477-2701 online

2018). However, one potential disadvantage is that learners must achieve all assessment outcomes to pass. For example, if a module consists of ten learning outcomes, each evaluated individually, learners must demonstrate an understanding of all ten outcomes – akin to achieving a perfect score of 100% in an examination. This assessment format is often open-book, so while students may have to address all assessment outcomes exhaustively, students can directly draw on information from textbooks, websites, and class notes. Nevertheless, many struggles with the analytical and evaluative skills required, and so success in criterion-referenced assessment requires continuous formative feedback; approaches such as the cognitive apprenticeship model (Lyons *et al.,* 2017) can be used to good effect here, maximising the likelihood of achieving assessment outcomes at the first attempt.

When reviewing assessment approaches used in pharmacy education, the Objective Structured Practical Examination (OSPE) (Ahmed *et al.,* 2011), which is a modification of the Objective Structured Clinical Examination (OSCE) pioneered by Harden and Cairncross (1980), is a third option that can be viewed as a compromise between the two previous examples. At its core, the OSPE assesses students by accounting for four factors:
1) Assessing the process and the product rather than assessing competency by viewing the final result; each step of the process is observed and assessed under controlled conditions;
2) Assessing breadth of skills: The skills required for professional practice can be specifically assessed by an OSPE activity, which is not always possible through a "*cookbook*" style practical work;
3) Student's approach to assessment: The approach to a task cannot be judged using typical assessment methods, yet this is an essential transferable and transversal skill. It can be assessed with OSPE through an overall performance rating;
4) Objectivity: Although mark schemes, rubrics, and other approaches ensure an acceptable level of objectivity during an assessment, different assessors will apply the same rubric slightly differently. In the OSPE, it is (usually) the same person who assesses all candidates for a particular assessment criteria.

From a student perspective, OSPEs have the advantage of being specific to the curriculum delivered at an institution (*vs* an externally set assessment), which allows for thorough preparation through formative feedback. As OSPEs have multiple assessors, the potential for biased judgements is reduced, and transparency is increased during the assessment. As multiple skills can be assessed at a single OSPE station, the process is efficient and reduces the overall examination time for candidates. From an institutional point of view, OSPEs provide valid and reliable assessment decisions that satisfy external stakeholders (Shirwaikar, 2015).

When designing an OSPE, considering the statistical framework in which the candidates' performance will be viewed is essential. For large cohorts, the borderline group method (Boursicot *et al*., 2007) has been used successfully as an alternative to the modified Angoff scheme. However, in smaller cohorts, the likelihood of having a sufficient number of borderline candidates to set the standard confidently is low; thus, the borderline regression model is preferred (Schoonheim-Klein *et al.,* 2009). In this approach, each OSPE station must have appropriately weighted item marks and an overall performance mark, e.g. fail, borderline, clear pass, good, and outstanding (Wood *et al*., 2006).

Educational setting and activity

The Diploma for Pharmacy Technicians is a two-year, competency-based programme mapped to the UK General Pharmaceutical Council's Initial Education & Training Standard for Pharmacy Technicians (Boughen & Fenn, 2020). It is assessed at Level 3 of the Regulated Qualifications Framework, representing a pre-degree-level qualification. In Belfast Metropolitan College, an integrated spiral curriculum is used where basic sciences are delivered and assessed during term one of the first year, after which students move to clinical pharmacy through a systems-oriented approach (Mawdsley & Willis, 2018).

Laboratory skills may be viewed as less influential in contemporary pharmacy technician training, given that extemporaneous dispensing is now considerably less prevalent in UK pharmacies, except in some hospital pharmacies and dedicated businesses (termed "*Specials manufacturers*"). Therefore, UK Pharmacy Technicians are unlikely to use practical, laboratory-like skills in their day-to-day practice. However, as many graduates seek employment in non-clinical roles (e.g. the pharmaceutical industry), it is essential to continue to embed core laboratory skills transferable to the compounding environment or preparation for more advanced roles (Burnett *et al*., 2003). These skills were historically delivered as a series of standard practical activities, with students working in pairs but submitting individual laboratory reports.

The development of the OSPE began by referring to the module assessment outcomes and extracting the core competencies that could be assessed via OSPE. Three core skills were selected, i.e. accurate weighing (OSPE 1), preparation of a dilution (OSPE 2), and calibration of a pH meter (OSPE 3), to be assessed in this format. These skills were selected as required for the pharmaceutics module in the second year (where extemporaneous dispensing skills are taught and assessed). The grading and characteristics of each global proficiency for each OSPE station were agreed upon by a quorum of six subject experts (two chemists, two practising pharmacists, and one academic pharmacist). First-year students from a sister Applied Science programme were approached to pilot this method; they had a similar background and level of experience and were divided into three groups. The first group undertook the legacy assessment (standard practical activities in pairs but individual laboratory reports (referred to as "*paired laboratory work*"), and the second group carried out the same standard practical activities but did so individually, submitting laboratory reports (referred to as "*individual laboratory work*"), and the third group undertook the proposed OSPEs. This approach helped refine the logistics of the process and account for confounding variables in the analysis (as explained later). Based on these findings, the move to OSPEs for the next academic year went ahead.

For the roll-out of the OSPEs, first-year pharmacy technician students were briefed about the assessment process and undertook three sessions of formative laboratory work through which they gained experience of the skills that would be summatively examined through OSPEs. For the examination period, the teaching laboratory was set up with three OSPE stations, each having one examiner who observed the candidates' performance and recorded their achievements on a score sheet. Candidates were provided with a set of instructions alongside any materials required. Before the examination, candidates were randomly divided into three groups, with each student assigned to start at a particular OSPE station. When the time elapsed (ten minutes per station), a buzzer sounded to signal the end of that OSPE, and candidates rotated to the next station.

Data collection and analysis

For the pilot study, Group 1 (legacy: paired laboratory work, individual reports; n = 12) and Group 2 (legacy: individual laboratory work, individual reports; n = 12) were graded using an established rubric that awarded an overall percentage mark. Group 3 (OSPE; n = 12) were given an overall percentage mark based on their performance at all three stations. As the sample size was small, Quade's analysis of covariance was selected to assess differences in individual performance, taking into account paired performance as a covariant.

For the main study, historical data from students who had undertaken the legacy assessment (paired laboratory work, individual reports; n = 63) were compared with that of the OSPE cohort (n = 65). Data were first transformed using the inverse hyperbolic sine function to satisfy the assumptions of the independent samples *t*-test. Learner engagement was evaluated through a short online questionnaire that reported results on a five-point Likert scale. The internal constancy of the Likert responses was assessed by Cronbach's alpha, and overall consensus scores for each question were calculated using a modification of Shannon entropy (Tastle & Wierman, 2007), which converts ordinal data into a dimensionless measure of dispersion (Eqn. 5 in reference).

All statistical analyses were performed using IBM SPSS Statistics for Windows, version 28.0.0.1.1 (IBM Corp., Armonk, N.Y., USA). Violin plots were constructed using GraphPad Prism for Windows, version 9.5.1 (GraphPad Software, San Diego, California USA). Written informed consent was obtained from all participants, in line with institutional ethics policy.

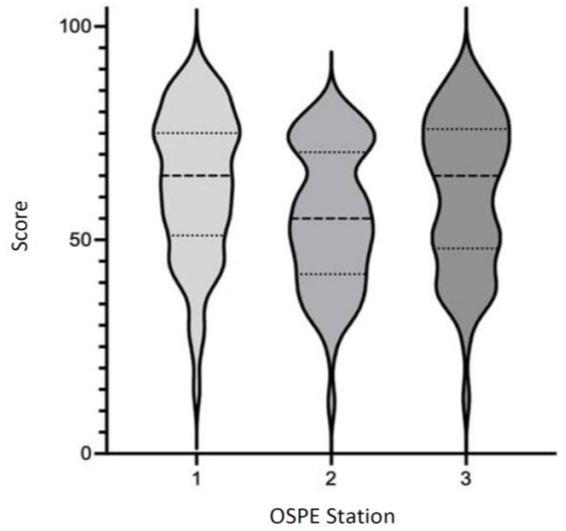
Evaluation

In the pilot study, the legacy assessment method was compared with the OSPE, considering the impact of working in pairs *vs* working alone. Analysis by Quade's ANCOVA revealed no significant difference between the groups (two-tailed *p*-value 0.504; *F* = 0.470), suggesting that the results obtained from the OSPE are in line with those obtained by the legacy assessment. Moreover, the move away from working in pairs to working alone does not appear to have any confounding impact. Table I summarises the performance of candidates in each of the OPSE stations.

Table I: Summary of OSPE performance

	Pilot study (n = 12)			Main study (n = 65)		
	OSPE 1	OSPE 2	OSPE 3	OSPE 1	OSPE 2	OSPE 3
Median (95% CI) (%)	61 ± 2.8	51 ± 3.5	69 ± 4.1	65 ± 4.2	55 ± 3.9	65 ± 4.3
Cut score (%)	44	41	45	43	44	46
Pass rate (%)	90	78	88	88	74	77

For the main study, candidate performance at each of the OSPE stations was largely homogeneous with scores following similar distributions (Figure 1).

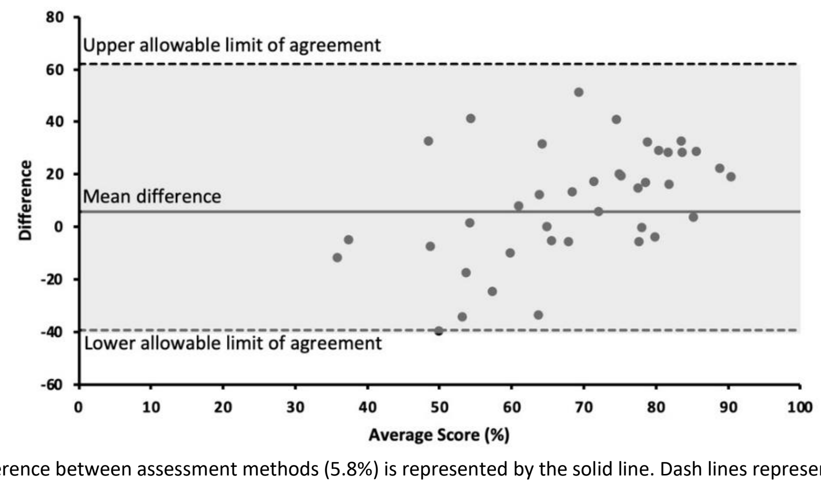


Dash line (---) represents the median score; Dot line (···) represents the 95% confidence intervals

Figure 1: Summary of scores for OSPE stations

The median score for the second OSPE (calculations plus practical manipulatives) was lower than the two other stations (only practical manipulatives), likely due to the increased complexity of this station compared with the two others. However, the use of a borderline regression model to establish the cut score means that the increased difficulty of this station was accounted for in assigning the pass/fail mark to candidates.

Regarding student performance, comparison of legacy *vs* OSPE groups revealed no significant difference between the two cohorts (two-tailed p-value 0.373; t = 0.896). Practical work was undertaken in pairs in the legacy cohort (although all students submitted individual reports), whereas the OSPE cohort undertook practical work individually. While a confounding impact from switching to individual practical work cannot be excluded, this switch appeared to have no significant effect in the pilot study, albeit at a lower statistical power. Hence, it can be inferred that OSPEs yield a similar spread of results compared to traditional assessment methods and do not significantly bias the pass rate (mean $\pm$ SD, OSPE *vs* legacy: 77 $\pm$ 19% *vs* 73 $\pm$ 21%). A Bland-Altman plot was constructed to highlight the impact of any outliers on this inference, where the difference in scores for the two assessment methods is on the y-axis and the average of the assessment scores on the x-axis (Figure 2). This method identified only a single borderline outlier, with all other student scores clustering about the mean and well within the limits of agreement ($\pm$ 1.96σ), confirming that the OSPEs are as valid an assessment instrument as the legacy approach.



The mean difference between assessment methods (5.8%) is represented by the solid line. Dash lines represent the upper and lower boundaries of agreement between the two assessment approaches

Figure 2: Bland-Altman plot of legacy *vs* OSPE performance

The results of the student evaluations showed strong internal consistency (Cronbach's α = 0.823), indicating a good level of reliability in these evaluations (Tavakol & Dennick, 2017). A strong consensus was found in the positive responses to Q1–3 and Q5, which is encouraging (Table II). Q4 (asking students if they felt the examiners' judgements were fair) yielded a lower consensus score, with two possible explanations.

Firstly, one of the examiners was a faculty member, unknown to the students, which could have potentially introduced an element of uncertainty/mistrust. Two quotes from students support this reason:

"what if they didn't see how much I weighed out"

"I don't think they could see the level [of liquid] from where they were standing"

Table II: Summary of student questionnaire responses

Question	Relative frequency*					C-score
	SA	A	N	D	SD	
The OSPEs were well-organised	0.43	0.53	0.03	0.03	0.00	0.78
I had enough time for each station	0.35	0.48	0.13	0.03	0.03	0.72
I felt prepared for the OSPEs	0.13	0.70	0.13	0.03	0.03	0.80
I feel that the judgements were fair	0.28	0.38	0.25	0.08	0.03	0.64
I prefer OSPEs to laboratory reports	0.75	0.18	0.08	0.00	0.00	0.80

*SA = strongly agree; A = agree; N = neutral; D = disagree; SD= strongly disagree

These concerns are valid, as in the most recent OSPEs, only faculty known to the students are used, and the observation process is over-emphasised by using statements like "*can I see the reading on the balance please*". A second explanation for the lower consensus score in Q4 is that students may not have fully understood the grading process. In the first iteration of the OSPEs, students were briefed about the grading process, but had not seen their record sheets (essentially the rubrics) against which they were being assessed. Some faculty view a rubric in the same light as a mark scheme for an exam and are resistant to sharing it with students before the assessment. However, allowing students to view rubrics helps them progress towards a goal and can thus be considered part of the formative assessment process (Jackson & Larkin, 2002). Hence, partially redacted candidate record sheets are made available to students before the OSPEs.

Conclusion

OSPEs are a means of securing authentic, criterion-referenced assessment decisions and can effectively replace traditional assessment methods. For this particular group of students, the early introduction of OSPEs has the added benefit of preparing them for OSCEs later in the program. This approach is also consistent with the upper levels of Miller's triangle, where pharmacy professionals are required to demonstrate "*shows how*" and "*does*" characteristics (McFadyen & Diack, 2017). The move to an OPSE approach was supported by the majority of the faculty, although some were cautious about the change, particularly regarding the time commitment involved. However, when comparing the effort-hours for marking laboratory reports and the time required for OSPEs, there was clear support for the latter. Other criticisms, such as stress to the student, role of the examiner, and fragmentation of the assessment, have been addressed elsewhere (Harden, 2015). Any change to the curriculum must have buy-in from all stakeholders – staff, students, and external groups such as employers. To address the latter, certificates of competency were issued, and students were guided in the development of a skills-based CV (résumé), providing employers with a more precise understanding of the skill set employees could possess rather than merely the title of a qualification.

Conflict of interest

The authors declare no conflict of interest.

Source of funding

The authors did not receive any funding.

References

Ahmed, K., Miskovic, D., Darzi, A., Athanasiou, T., & Hanna, G. B. (2011). Observational tools for assessment of procedural skills: A Systematic Review. *The American Journal of Surgery,* **202**(4), 469–480. https://doi.org/10.1016/j.amjsurg.2010.10.020

Boughen, M., & Fenn, T. (2020). Practice, skill mix, and education: The evolving role of pharmacy technicians in

Great Britain. *Pharmacy,* **8**(2), 50. https://doi.org/10.3390/pharmacy8020050

Boursicot, K. A. M., Roberts, T. E., & Pell, G. (2007). Using borderline methods to compare passing standards for OSCEs at graduation across three medical schools. *Medical Education,* **41**, 1024–1031. https://doi.org/10.1111/j.1365-2923.2007.02857.x

Burnett, D., Dooley, M. J., & Wall, D. (2003). Cell-based therapies—A role for pharmacy technicians. *Journal of Pharmacy Practice and Research,* **33**, 296–298 https://doi.org/10.1002/jppr2003334296

Gericke, N., Högström, P., & Wallin, J. (2022). A systematic review of research on laboratory work in secondary school. *Studies in Science Education,* **59**(2), 245–285. https://doi.org/10.1080/03057267.2022.2090125

Gott, R., & Duggan, S. (2002). Problems with the assessment of performance in practical Science: Which way now? *Cambridge Journal of Education,* **32**(2), 183–201. https://doi.org/10.1080/03057640220147540

Harden, R. M. (2015). Misconceptions and the OSCE. *Medical Teacher,* **37**(7), 608–610. https://doi.org/10.3109/0142159X.2015.1042443

Harden, R. M., & Cairncross, R. G. (1980). Assessment of practical skills: The objective structured practical examination (OSPE). *Studies in Higher Education,* **5**(2), 187–196. https://doi.org/10.1080/03075078012331377216

Jackson, C. W., & Larkin, M. J. (2002). Teaching students to use grading rubrics. *Teaching Exceptional Children*, **35**(1), 40–45. https://doi.org/10.1177/004005990203500106

Lee, H. (2023). The rise of ChatGPT: Exploring its potential in medical education. *Anatomical Sciences Education,* **00**, 1–6. https://doi.org/10.1002/ase.2270

Lyons, K., McLaughlin, J. E., Khanova, J., & Roth, M. T. (2017). Cognitive apprenticeship in health sciences education: A qualitative review. *Advances In Health Science Education,* **22**, 713–739. https://doi.org/10.1007/s10459-016-9707-4

Mawdsley, A., & Willis, S. (2018). Exploring an integrated curriculum in pharmacy: Educators' perspectives. *Currents in Pharmacy Teaching & Learning,* **10**(3), 373–381 https://doi.org/10.1016/j.cptl.2017.11.017

McFadyen, M. C. E., & Diack L. (2017). I can step outside my comfort zone. *Pharmacy,* **5**(4), 59. https://doi.org/10.3390/pharmacy5040059

Newton, P. E. (2011). A level pass rates and the enduring myth of norm-referencing. *Research Matters: A Cambridge Assessment publication,* Special Issue **2**, 20–26. https://www.cambridgeassessment.org.uk/Images/567580-a-level-pass-rates-and-the-enduring-myth-of-norm-referencing.pdf

Pereira, A. G., Woods, M., Olson, A. P., Van Den Hoogenhof, S., Duffy, B. L., & Englander, R. (2018). Criterion-based assessment in a norm-based world: How can we move past grades? *Academic Medicine,* **93**(4), 560–564. https://doi.org/10.1097/ACM.0000000000001939

Schoonheim-Klein, M., Muijtjens, A., Habets, L., Manogue, M., Van Der Vleuten, C., & Van Der Velden, U. (2009). Who will pass the dental OSCE? Comparison of the Angoff and the borderline regression standard setting methods. *European Journal of Dental Education,* **13**, 162–171. https://doi.org/10.1111/j.1600-0579.2008.00568.x

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

Sosabowski, M. H., & Gard, P. R. (2008). Pharmacy education in the United Kingdom. *American Journal of Pharmaceutical Education,* **72**(6), 130. https://doi.org/10.5688/aj7206130

Tastle, W. J., & Wierman, M. J. (2007). Consensus and dissention: A measure of ordinal dispersion. *International Journal of Approximate Reasoning,* **45**(3), 531–545. https://doi.org/10.1016/j.ijar.2006.06.024

Tavakol, M., & Dennick, R. (2011). *Making sense of cronbach's alpha. International Journal of Medical Education,* **2**, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Wood, T. J., Humphrey-Murto, S. M., & Norman, G. R. (2006). Standard setting in a small scale OSCE: A comparison of the modified borderline-group method and the borderline regression method. *Advances in Health Science Education,* **11**, 115–122. https://doi.org/10.1007/s10459-005-7853-1