RESEARCH

Reliability and Validity of a Checklist to Evaluate Student Performance in a Problem-Based Learning Group

Francine D. Salinitri, PharmD,^a Alison M. Lobkovich, PharmD,^b Brian L. Crabtree, PharmD,^c Sheila M. Wilhelm, PharmD^a

^a Eugene Applebaum College of Pharmacy and Health Sciences, Wayne State University, Detroit, Michigan

^b Harper University Hospital, Detroit, Michigan

^c Mercer University College of Pharmacy, Atlanta, Georgia

Submitted January 21, 2018; accepted September 23, 2018; published October 2019.

Objective. To validate a problem-based learning (PBL) evaluation checklist to assess individual Doctor of Pharmacy (PharmD) students' performance in a group.

Methods. In 2013, a performance checklist was developed and standardized. To evaluate the reliability and discriminant validity of the checklist, pharmacy students' evaluation scores from 2015-2016 were assessed along with overall program grade point averages (GPA), and scores on knowledge and problem-solving examinations. Predictive analysis software was used to analyze the data.

Results. Seventy facilitators generated 1506 evaluation reports for 191 (90 third-year and 101 second-year) students over eight PBL cases. The mean (SD) total score was 40.6 (2.5) for P3s and 39.1 (2.7) for P2s out of a possible 44.2 points. Students' scores improved each semester. Interrater reliability based on intraclass correlation coefficient for all cases was 0.67. Internal reliability as determined by Cronbach alpha was >0.7 for all binary checklist items across all cases. Discriminant validity assessed using Pearson correlation coefficient showed that the total score from the checklist did not correlate with knowledge or problem-solving examination scores.

Conclusion. This unique PBL checklist proved to be a reliable and valid tool to assess student performance in small group sessions in a PharmD curriculum.

Keywords: problem-based learning, validity, reliability, checklist, assessment

INTRODUCTION

Problem-based learning (PBL) is a common pedagogy used in higher education, specifically in health care professions' education.^{1,2} The majority of Doctor of Pharmacy (PharmD) programs report using PBL within their curricula.³ Problem-based learning is the student-directed pedagogy that most closely mimics clinical practice in a classroom setting. Skills developed through PBL include problem-solving, critical-thinking, clinical-reasoning, self-directed learning, collaborative practice, flexible knowledge, and intrinsic motivation.⁴⁻⁶ In order to acquire these skills, the PBL process should be authentically implemented using small, collaborative groups of students guided by a trained facilitator.⁷ The stimulus for group learning is an ill-structured, complex, open-ended problem, such as a realistic patient case.⁴

At Wayne State University (WSU), we integrated PBL within our pharmacotherapy modules in 2006. In 2014, we separated PBL from the modules to create a PBL four-course series called Pharmacotherapeutic Problem Solving, which spans the second- and third-professional (P2 and P3) years. Problem-based learning in a pharmacy curriculum offers a student-directed approach to learning pathophysiology, pharmacology, medicinal chemistry, and pharmacotherapeutics within the context of a patient case. The content learned in the PBL courses is unique but complimentary to the content of the concurrently delivered didactic pharmacotherapeutic module courses. The framework for the PBL courses was developed by aligning PBL skills with specific program-level ability-based outcomes.

Assessment of knowledge acquired from PBL experiences has been extensively described in the literature. In pharmacy education literature, the use of written case-based examinations at the end of the PBL process⁸⁻¹⁰ or as a pre- and post-test assessment of knowledge^{11,12} have been reported. In addition to summative examinations,

Corresponding Author: Francine D. Salinitri, Eugene Applebaum College of Pharmacy and Health Sciences, Wayne State University, 259 Mack Ave., Detroit, MI 48201. Tel: 313-577-0453. Fax: 313-577-5369. E-mail: fsalinitri@wayne.edu

weekly, written, patient-care plan submissions have been evaluated using a tick-box rubric.⁸ Others report using Likert scale rubrics to assess written submissions encompassing the entire PBL process, including the development of facts, hypotheses, and learning questions, as well as drug-related problems and their resolution.¹³ In the medical literature, students' written patient summary statements have been assessed using rubrics based on Likert-scale ratings.¹⁴

A variety of assessment tools evaluating skills developed by students during PBL have also been described. Summative evaluations of PBL skills have been reported using objective structured clinical examinations (OSCEs)¹⁵⁻¹⁷ or problem-solving written examinations.⁹ Peer evaluations of performance within PBL sessions have been used,^{8,9,18,19} as have student self-evaluations.^{9,11,20} The literature that exists surrounding facilitator evaluation of student performance within a small group is limited. Ross and colleagues report on the use of a facilitator evaluation that encompasses PBL skills, such as participation, cooperation, and communication skills.⁹ Romero and colleagues describe the use of verbal feedback for students' performance with a PBL group.¹⁰ Sim and colleagues discuss and evaluate the use of a five-point Likert-scale facilitator evaluation instrument.²¹ This tool assessed participation and communication skills, cooperation or team-building skills, comprehension or reasoning skills, and knowledge or information-gathering skills. Facilitators were surveyed, and 88% agreed that the tool was easy to use (n=34). Interrater reliability was assessed using an intraclass correlation coefficient that indicated overall consistency among facilitators' scores with some strict (4 of 34) or indiscriminant (8 of 34) graders.²¹

Over the past decade at WSU, we used facilitator-graded Likert scales to evaluate student performance within the small group setting. Each facilitator evaluated the members of their group, which consisted of 8 to 9 students. Anecdotally, we observed that facilitators did not discriminate between students' performance using this assessment tool, leading to inflated evaluation scores. In order to address the high facilitator scores, we instituted tally sheets within the sessions, which were used by both the facilitator and two student members of the group to determine the quantity and quality of comments made by each student. The facilitator was asked to use these tally sheets to inform their Likert scale-based rating of each student in the group. However, feedback from facilitators and students regarding the tally sheets indicated that they inhibited the students' ability to participate and the facilitators' ability to coach the group. In addition, student scores still appeared to be inflated.

At WSU, we employ a multi-modal assessment of student learning and performance within the PBL course series using four types of assessments that occur multiple times throughout the semester (Figure 1). One of the assessments requires students to develop individual patient care plans for each of the four cases during the self-directed learning process. These care plans are graded by the case writers, who are content experts. The plans are submitted and graded electronically using rubrics. To maintain the student-driven authenticity of the PBL pedagogy, rubrics for providing students with general feedback regarding their individual patient care plans were developed rather than providing students with content-specific answers. Instead, content-specific answers for the plans are discussed and consensus is reached within the small group during subsequent facilitated sessions. In addition, midterm and final examinations assessing pharmacotherapeutic and disease state knowledge acquisition are administered, each of which covered two patient cases. Written problem-solving examinations were administered with the midterm and final knowledge-based assessments. During the problem-solving examination, the students were presented with a patient case involving a disease state that had not been formally taught within the curriculum. The students were provided with selected primary, secondary, and tertiary resources regarding the disease state. The students individually completed the steps they would ordinarily complete with their group: identifying facts from the case, developing related hypotheses and learning questions, answering those learning questions using the provided resources, critiquing the provided resources, and developing an answer to the problem in the case through responding to a variety of question formats (eg, multiple choice, multiple answer, short answer, matching, essay).

We developed a fourth type of assessment, which is a novel checklist-based tool for facilitators to use within the small-group sessions to evaluate student performance (Appendix 1). The tool was developed to evaluate the following course objectives, which were derived from the Center for the Advancement of Pharmacy Education 2013 Educational Outcomes²²: differentiate relevant patient characteristics; generate hypotheses or learning questions; discuss clinical and scientific principles of the disease state; discuss clinical and scientific principles of the pharmacotherapies; identify drug-related problems (DRPs) through a systematic evaluation process; formulate a patient-specific, literature-supported patient plan to resolve DRPs; employ an effective strategy to identify relevant scientific literature and resources; and evaluate the scientific literature to formulate a solution to the problem. Facilitators used the tool over the span of three group

Figure 1. Wayne State University Problem-Based Learning (PBL) Course Series Assessment Blueprint for the P2 and P3 2015-16 Year

sessions, at the end of which they also provided a global assessment of student professionalism and performance. The tool was completed for each patient case (four times) over the course of the semester for each student. The checklist was developed as a unique tool to assess problem-solving skills. We hypothesized that this binary checklist was a reliable and internally valid tool to measure student performance within a PBL small group. In addition, we hypothesized that the facilitator checklist provided discriminant validity for student performance and, therefore, would be weakly correlated with other measures of performance including PBL-course knowledge-based and problem-solving examinations, and program grade point average (GPA). The goal of this study was to evaluate the reliability and discriminant validity of the checklist evaluation tool for assessing pharmacy students' performance within a small-group PBL environment.

METHODS

A standardized facilitator evaluation checklist for the assessment of student performance in a PBL small group was developed in 2013. The checklist is provided in Appendix 1. The PBL planning committee developed a list of eight course objectives derived from the CAPE

2013 Educational Outcomes²² and WSU's program ability-based outcomes. For each objective, the committee identified tasks and skills that would measure achievement of the objective. Through this process, there were 52 items identified for the checklist. Based on our curricular experience with previous facilitator evaluation forms that used Likert scales, and the concern for associated grade inflation, we chose to employ a binary checklist approach where a facilitator would give each student a check mark for each item that was demonstrated in one of the three facilitated group sessions. By using a checklist approach, we hoped to achieve a more objective measure of observed student performance within a small-group PBL experience. Students only received the check mark on the first instance that they demonstrated a specific task. Subsequent instances were taken into account by the facilitator in the global evaluation questions that applied to all three sessions. This global assessment allowed the facilitator to differentiate students who consistently engaged in the PBL process from those who did not. The specific questions used are found in the Overall Professional Performance and Global Assessment sections of Appendix 1.

Once the checklist was finalized and the PBL committee had reviewed all of the items and agreed on the final

language of the checklist, the Angoff Method for Standard Setting was used to determine the weighting of each item on the checklist.²³ In order to employ this process, we identified 10 judges²⁴ who were familiar with our program and who work with and hire our graduates. These judges were pharmacy practice and pharmaceutical sciences faculty members and included PBL committee members, volunteer preceptors, and area pharmacy administrators from a variety of practice settings. The judges were provided with the checklist and asked to determine what percent of the time a minimally competent pharmacy student at the end of their third professional year would achieve each task. A minimally competent student was defined as someone who would be able to perform with a passing score of 70% within our program. The estimates for each item provided by the judges were then averaged, and those averages were used to determine the weighted score for each item.

The checklist was used for one academic year to assess both P2 and P3 students. The PBL planning committee then evaluated the checklist based on student scores and facilitator feedback and determined that some items were very similar and could be combined. Other items were assessed elsewhere in the course and were eliminated from the checklist. The revised checklist measured 37 items and eight objectives and has been used since fall 2015.

In order to determine the reliability and discriminant validity of the evaluation checklist, we performed a retrospective analysis of facilitator evaluations of P2 and P3 students from the 2015-2016 academic year. Data reports

were generated and compiled from E*VALUE (MedHub, Minneapolis, MN) as well as from reports of student program GPAs and knowledge and problem-solving examination performance in each PBL course. Other data included in the database was student year (P2 or P3), PBL group assignment, facilitator assignment, and facilitator type. The facilitators responsible for evaluating student performance were volunteers, first-year pharmacy residents, and Pharmaceutical Science and Pharmacy Practice faculty at WSU. The facilitators were trained for facilitation and evaluation through a structured training program, which has been described elsewhere.²⁵ All data were de-identified and numerically coded. This study was approved by the WSU Institutional Review Board as a Quality Improvement project.

Descriptive statistics were used to describe our student population and mean evaluation scores. Student *t* test was used to determine the similarity between standardized judges' scores on the checklist items versus scores achieved by the students during their PBL sessions. Interrater reliability of all facilitator checklist assessments during the 2015-2016 academic year for P2 and P3 students was assessed using the intraclass correlation coefficient (ICC). Cronbach alpha was used to assess the reliability of the items within the checklist for each case for P2 and P3 courses. A conventional cut-off of 0.7 or higher was defined as reliable.²⁶ Discriminant validity of the checklist was assessed using the Pearson correlation coefficient (PCC). The PCC was used to assess correlation between

Figure 2. Mean Problem-Based Learning Checklist Performance Scores Out of a Maximum of 44.2 Points for Each Case for Both P2 and P3 Classes

facilitator evaluation scores, overall program GPA, knowledge examination scores, and problem-solving examination scores. A p value $\leq .05$ (two-tailed) was considered significant. The PCC was also used to assess correlation between items on the checklist that were achieved by $<80\%$ of students. A PCC of less than 0.3 was considered weakly correlated, while 0.3-0.49 was considered moderately correlated, and greater than or equal to 0.5 was considered strongly correlated.²⁷ The IBM Statistical Package for the Social Sciences (SPSS, Armonk, NY) predictive analytics software was used for statistical analyses.

RESULTS

In 2015-2016, P2 and P3 students enrolled in WSU's PBL courses completed 16 cases. These courses enrolled 191 students, which were divided into 24 groups and facilitated by 70 facilitators who generated 1506 evaluation reports (Figure 1). The mean checklist score for the P2 class over eight evaluations was 39.1 (SD=2.7) out of 44.2 possible points that students could achieve for their performance during the small group PBL sessions. The mean checklist score for the P3 class was 40.6 (SD=2.5). Both P2 and P3 mean evaluation scores improved over the course of each semester for the majority of cases (Figure 2).

The ICC as a measure of interrater reliability was 0.67 (95% CI=0.57-0.73) for all P2 and P3 checklist scores. Cronbach alpha as a measure of internal reliability of the checklist was computed for each of the 16 cases and ranged from 0.57 to 0.72 for P2 cases and 0.60 to 0.77 for P3 cases (Table 1). When checklist item number 47 (Appendix 1), "What overall grade on a scale of 0-100% would you give this student?" was omitted from the analysis, the Cronbach alpha increased from 0.79 to 0.92 for P2 cases and from 0.79 to 0.95 for P3 cases.

To assess the internal validity of the performance checklist, items achieved less frequently by students were analyzed as a subset of the checklist. All items achieved by fewer than 80% of students were positively correlated with one another. Within the P2 class, 13 checklist items that were achieved by fewer than 80% of students resulted in 78 pairwise correlations, and 73 of the 78 correlations were significant ($p \leq .05$). Within these 73 significant correlations, there were 56 pairwise comparisons that were moderately or strongly correlated ($PCC > 0.3$) (Table 2). Within the P3 class, five checklist items were achieved by fewer than 80% of students, resulting in 10 pairwise correlations, all of which were significantly correlated ($p \leq .05$). Nine of the 10 pairwise correlations were moderately or strongly correlated ($PCC > 0.3$) (Table 3).

The PBL checklist evaluation scores from facilitators on students' performance in small group session were compared with other measures of performance within the

Table 1. Internal Reliability Scores for a Checklist Used to Evaluate P2 and P3 Students' Performance on Problem-based Learning Cases

Year	Semester and PBL Case No.	Cronbach Alpha	Cronbach Alpha With Item Omitted ^a
P2	Fall 2015		
	1	0.66	0.92
	2	0.67	0.87
	3	0.65	0.79
	4	0.63	0.85
	Winter 2016		
	1	0.72	0.90
	2	0.66	0.87
P3	Fall 2015		
	1	0.76	0.93
	2	0.67	0.87
	3	0.70	0.93
	4	0.75	0.93
	Winter 2016		
	1	0.69	0.91
	2	0.73	0.92
	3	0.60	0.79
	4	0.77	0.95

^a Data were analyzed with the removal of item 47 of the checklist: "What overall grade on a scale of 0-100% would you give this student?" Item 47 was the only non-binary, subjective item on the checklist

When the objective, binary items that remained were analyzed, the Cronbach's alpha rose to >0.7 for all cases assessed
P2=second professional year students, P3=third professional year students, PBL=problem-based learning

PBL courses (knowledge and problem-solving examination scores) and overall program performance as measured by program GPA to assess discriminant validity of the checklist. With regard to P2 student performance, checklist evaluation scores for only three of the eight cases correlated significantly, albeit weakly ($PCC < 0.3$), with student performance on the knowledge examinations. These included case 4 of the fall semester and cases 1 and 2 of the winter semester (Table 4). None of the evaluation scores during the fall or winter semester were significantly correlated with student performance on the problem-solving examinations. When correlated with overall program GPA, the evaluation scores for only two of the eight cases were significantly correlated (case 1 and case 2 of the winter semester, $PCC < 0.3$) (Table 5).

With regard to P3 student performance, evaluation scores from two of the eight cases (cases 2 and 3 of the

Table 2. Checklist Item Internal Correlation for P2 Student Performance in Small Group Sessions^a

Checklist Item No. (Appendix 1)	10	13	15	17	20	21	29	30	32	33	34	36	37
10	1												
13	.26 ^b	1											
15	.63 ^b	.08	1										
17	.42 ^b	.41 ^b	.35 ^b	1									
20	.61 ^b	.44 ^b	.50 ^b	.60 ^b	1								
21	.39 ^b	.55 ^b	.38 ^b	.60 ^b	.59 ^b	1							
29	.52 ^b	.39 ^b	.58 ^b	.57 ^b	.59 ^b	.53 ^b	1						
30	.34 ^b	.24 ^b	.29 ^b	.43 ^b	.23 ^b	.37 ^b	.45 ^b	1					
32	.20 ^b	.27 ^b	.23 ^b	.24 ^b	.22 ^b	.41 ^b	.32 ^b	.40 ^b	1				
33	.33 ^b	.29 ^b	.28 ^b	.30 ^b	.16	.39 ^b	.42 ^b	.52 ^b	.73 ^b	1			
34	.27 ^b	.24 ^b	.23 ^b	.14	.17	.35 ^b	.33 ^b	.42 ^b	.70 ^b	.86 ^b	1		
36	.21 ^b	.33 ^b	.40 ^b	.24 ^b	.38 ^b	.50 ^b	.40 ^b	.15	.51 ^b	.55 ^b	.56 ^b	1	
37	.41 ^b	.39 ^b	.49 ^b	.36 ^b	.47 ^b	.63 ^b	.53 ^b	.36 ^b	.58 ^b	.65 ^b	.66 ^b	.71 ^b	1

^a Checklist item internal correlation for P2 student performance in small group sessions is reported as Pearson correlation coefficients (PCC). Correlation of the 13 checklist items that were achieved by fewer than 80% of P2 students (n=101) resulted in 78 pairwise correlations. 73 of the 78 correlations were significant ($p \leq .05$), and 56 of the pairs were moderately or strongly correlated, defined as a $PCC > 0.3$. See Appendix 1 for details regarding individual checklist item numbers

^b Indicates $p \leq .05$

winter semester) were significantly but weakly correlated ($PCC < 0.3$) with student performance on the knowledge examinations. Weak correlations ($PCC < 0.3$) were found between the problem-solving examinations and evaluation scores from cases 1, 2, and 3 of the winter semester (Table 4). Lastly, students' overall program GPA was weakly correlated ($PCC < 0.3$) with checklist evaluation scores from case 3 of the fall semester and cases 2 and 3 of the winter semester (Table 5).

DISCUSSION

Assessment of skills gained during the PBL process is necessary to determine whether educational outcomes have been met. The standards for effective and efficient

performance evaluation within a PBL small group are not well described or studied. Our results describe the assessment of the reliability and validity of a checklist-based tool that can be used for the evaluation of student performance within a PBL small group in a pharmacy curriculum.

In small group PBL sessions, facilitators use the checklist to evaluate students' performance by indicating whether or not each task was performed at least once. Students could achieve a maximum total of 44.2 points over the three facilitated group sessions. Mean checklist evaluation scores were higher in the P3 class compared with the P2 class. This may have been because the P3 students had had an additional year of content knowledge and experience within the PBL process and were more familiar with the expected performance items on the checklist. Mean evaluation scores for both P2 and P3 students generally increased over the course of each semester (Figure 2). Within the P2 course, there was a dip in mean score from fall case 2 to fall case 3. This may have been because of the narrow scope of the case content, which did not provide enough opportunities for all of the students within the group to share information and achieve the specific items on the checklist. For example, therapies for the case disease state were limited thus, some students may not have had the opportunity to discuss clinical and scientific principles of the pharmacotherapies.

Interestingly, the first case in both the fall and winter semesters resulted in the lowest scores of the year, as though performance reset to a baseline level with the start

Table 3. Checklist Item Internal Correlation for P3 Student Performance in Small Group Sessions^a

Checklist Item No. (Appendix 1)	10	17	20	29	36
10	1				
17	.40 ^b	1			
20	.62 ^b	.61 ^b	1		
29	.52 ^b	.61 ^b	.62 ^b	1	
36	.25 ^b	.31 ^b	.37 ^b	.41 ^b	1

^a Correlation of the five checklist items that were achieved by fewer than 80% of P3 students (n=98) resulted in 10 pairwise correlations that were all significant ($p \leq .05$). Nine of the 10 pairwise correlations were moderately or strongly correlated (Pearson Correlation Coefficient (PCC) > 0.3). See Appendix 1 for details regarding individual checklist item numbers

^b Indicates $p \leq .05$

	Fall 2015						Winter 2016										
	Case 1		Case 2		Case 3		Case 4		Case 1 Winter		Case 2 Winter		Case 3 Winter		Case 4 Winter		
	P2	P3	P2	P3	P2	P3	P2	P3	P2	P3	P2	P3	P2	P3	P2	P3	
Type of Assessment	N=98	N=89	N=98	N=89	N=100	N=89	N=100	N=89	N=89	N=99	N=88	N=98	N=88	N=99	N=88	N=99	N=88
Knowledge examination	-.03	.17	.14	.14	.16	.08	.27 ^b	.08	.12	.22 ^b	.14	.22 ^b	.25 ^b	-.03	.24 ^b	.12	.10
Problem-solving examination	-.08	.04	-.05	.05	.09	.05	.10	-.14	.12	-.35 ^b	.19 ^b	-.31 ^b	.22 ^b	-.07	.15		

P2=second professional year, P3=third professional year

To assess the internal validity of the checklist, we evaluated whether the performance checklist was able to differentiate performance among students by assessing correlation between items. Items achieved by fewer students (less than 80%) were analyzed. In both the P2 and P3 courses, more than 75% of the items that were achieved by less than 80% of students were either moderately or strongly correlated with one another, illustrating that

Table 5. P2 and P3 Evaluation Scores vs Overall Program Grade Point Average

Case No. and Semester	P2, N	Pearson Correlation	p Value	P3, N	Pearson Correlation	p Value
Case 1 fall	98	.03	.76	90	.19	.07
Case 2 fall	98	.08	.46	90	.19	.08
Case 3 fall	100	.18	.07	89	.24	.02 ^a
Case 4 fall	100	.18	.07	89	.19	.08
Case 1 winter	100	.20	.05 ^a	88	.18	.09
Case 2 winter	99	.26	.01 ^a	88	.29	.01 ^a
Case 3 winter	100	.19	.07	88	.23	.03 ^a
Case 4 winter	100	.13	.19	88	.08	.46

^a Indicates $p \leq .05$

P2 N=number of second professional year students

P3 N=number of third professional year students

stronger performers tended to achieve these items. Of the items identified, 10 of the 13 items that were achieved by fewer than 80% of P2 students were predicted to be performed by 80% or fewer students based on the judges' standard setting scores. Among P3 students, the five items that were achieved by fewer than 80% of students were a subset of those same items identified in the P2 class. This indicates that students' PBL skills continue to develop while moving through the PBL course series from the P2 to the P3 year. We explored potential explanations for why P3 students did not fully achieve the five identified items. For the items regarding the pathophysiology of the disease state, pharmacokinetic and pharmacodynamic parameters, and drug cost and pharmacoeconomic considerations, we identified that lack of case complexity (ie, fewer medications and drug related problems included in the case) did not afford multiple students the opportunity to address these issues in their group discussion. We have since reviewed PBL cases across the course series with the goal of increasing their complexity to allow ample opportunities for students to share their knowledge within the group setting in these subject areas. The lack of achievement of the items regarding sharing and defending literature search terms and strategies, and discussing the limitations and sources of bias of supporting literature may have been due to time constraints within the group setting, and the subsequent lack of prompting from the facilitator. Therefore, we have developed modified PBL courses for P3 students that emphasize critical evaluation, application, and synthesis of supporting literature.

The performance checklist was intended to measure students' problem-solving skills during each PBL session. Checklist scores were compared against knowledge-based and problem-solving examinations, and program GPA to assess discriminant validity of the checklist. As hypothesized, our results indicated that the checklist scores were weakly correlated or not correlated

with knowledge-based examinations, problem-solving examinations, or program GPA. The problem-solving examinations were developed with the intent of evaluating students' problem-solving skills gained through the small group PBL experience. However, the examinations were individual, written examinations that did not afford the students the opportunity to demonstrate the skills assessed using the checklist. Potentially, a performance-based examination would provide opportunities for summative evaluation of these types of skills. To clarify, our P2 and P3 classes were two different groups of students rather than the same cohort of students followed from their second to third year in pharmacy school. In future research, we may consider evaluating the same group of students as they move through their P2 and P3 years, which would allow us to control for potential confounding variables that inevitably occur between two different classes of students.

CONCLUSION

To our knowledge, this is the first checklist-based instrument that has demonstrated reliability and discriminant validity in evaluating student performance in a small group PBL environment. Results of this study support that our checklist to evaluate student performance within a PBL small group has the potential to be generalized for use across any pharmacy PBL curriculum. The checklist demonstrated interrater reliability among a variety of facilitators including pharmaceutical science faculty members, pharmacy practice faculty members, volunteer adjunct faculty members from various practice settings, and pharmacy residents. Variations that do exist between facilitators could be addressed through continued improvements in standardizing facilitator training. Results of the Cronbach alpha could also be generalized to other programs to help identify cases that require improvement in order to enhance the PBL experience. For instance, the

cases identified as having low alpha scores were replaced with disease state topics that had more complexity. Finally, this checklist differentiates student performance and offers pharmacy programs a mechanism for evaluating skills that are not routinely assessed through conventional examination-based assessments.

REFERENCES

1. Ge X, Planas LG, Huang K. Guest editors' introduction: special issue on problem-based learning in health professions education/ toward advancement of problem-based learning research and practice in health professions education: motivating learners, facilitating processes, and supporting with technology. *Interdisciplinary J Prob Based Learn*. 2015;9(1):Article 5.
2. L'Ecuyer KM, Pole D, Leander SA. The use of PBL in an interprofessional education course for health care professional students. *Interdisciplinary J Prob Based Learn*. 2015;9(1):Article 6.
3. Stewart DW, Brown SD, Clavier CW, Wyatt J. Active-learning processes used in US pharmacy education. *Am J Pharm Educ*. 2011;75(4):Article 68.
4. Barrows HS. A taxonomy of problem-based learning methods. *Med Educ*. 1986;20:481-486.
5. Barrows HS. *Practice-Based Learning: Problem-Based Learning Applied to Medical Education*. Springfield, IL: Southern Illinois University School of Medicine. 1994;145.
6. Hmelo-Silver CE. Problem-based learning: what and how do students learn. *Educ Psychol Rev*. 2004;16:235-266.
7. Hmelo-Silver CE, Barrows HS. Goals and strategies of a problem-based learning facilitator. *Interdisciplinary J Prob Based Learn*. 2006;1(1):21-39.
8. Strohfeldt K, Grant DT. A model for self-directed problem-based learning for renal therapeutics. *Am J Pharm Educ*. 2010;74(9):Article 173.
9. Ross LA, Crabtree BL, Theilman GD, Ross BS, Cleary JD, Byrd HJ. Implementation and refinement of a problem-based learning model: a ten-year experience. *Am J Pharm Educ*. 2007;71(1):Article 17.
10. Romero RM, Eriksen SP, Haworth IS. A decade of teaching pharmaceuticals using case studies and problem-based learning. *Am J Pharm Educ*. 2004;68(2):Article 31.
11. Juba KM, Ricca BP. Design of a problem-based learning pain and palliative care elective course. *Curr Pharm Teach Learn*. 2014;6(3):421-428.
12. Webster AA, Riggs RM. A quantitative assessment of a medicinal chemistry problem-based learning sequence. *Am J Pharm Educ*. 2006;70(4):Article 89.
13. Dolder CR, Olin JL, Alston GL. Prospective measurement of a problem-based learning course sequence. *Am J Pharm Educ*. 2012;76(9):Article 179.
14. Smith S, Kogan JR, Berman NB, Dell MS, Brock DM, Robins LS. The development and preliminary validation of a rubric to assess medical students' written summary statements in virtual patient cases. *Acad Med*. 2016;91(1):94-100.
15. Des Marchais JE, Vu NV. Developing and evaluating the student assessment system in the preclinical problem-based curriculum at Sherbrooke. *Acad Med*. 1996;71(3):274-283.
16. Goodyear HM. Problem based learning in a junior doctor teaching programme. *Arch Dis Child*. 2005;90(3):275-278.
17. Salinitri FD, O'Connell MB, Garwood CL, Lehr VT, Abdallah K. An objective structured clinical examination to assess problem-based learning. *Am J Pharm Educ*. 2012;76(3):Article 44.
18. Roberts C, Jorm C, Gentilcore S, Crossley J. Peer assessment of professional behaviours in problem-based learning groups. *Med Educ*. 2017;51(4):390-400.
19. Kritikos VS, Woulfe J, Sukkar MB, Saini B. Intergroup peer assessment in problem-based learning tutorials for undergraduate pharmacy students. *Am J Pharm Educ*. 2011;75(4):Article 73.
20. Khoiriyah U, Roberts C, Jorm C, Van der Vleuten CP. Enhancing students' learning in problem based learning: validation of a self-assessment scale for active learning and critical thinking. *BMC Med Educ*. 2015;15(1):Article 140.
21. Sim NH, Azila SM, Lian NMA, Tan LH. A simple instrument for the assessment of student performance in problem-based learning tutorials. *Ann Dent Univ Malaya*. 2006;35(9):634-641.
22. Medina MS, Plaza CM, Stowe CD, et al. Center for the Advancement of Pharmacy Education 2013 educational outcomes. *Am J Pharm Educ*. 2013;77(8):Article 162.
23. Angoff WH. *Scales, Norms, and Scores Equivalent: Educational Testing Service*. Princeton, NJ; 1984:8-9.
24. Jaeger RM. Selection of judges for standard-setting. *Education Measurement: Issues and Practices*. 1991;10(2):3-14.
25. Salinitri FD, Wilhelm SM, Crabtree BL. Facilitating facilitators: enhancing PBL through a structured facilitator development program. *Interdisciplinary J Prob Based Learn*. 2015;9(1):Article 11.
26. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess*. 1994;6(4):284-290.
27. Evans JD. *Straightforward Statistics for the Behavioral Sciences*. Pacific Grove, CA; Brooks/Cole Publishing Company. 1996.

Appendix 1. Problem Based Learning (PBL) Facilitator Checklist to Evaluate Student Performance within Small Group Sessions. Each Checklist Item Represents Possible Tasks that a Student Could Perform in a Session.

Course Objective and Number (O#)	Checklist Item Number
O1. Differentiate relevant patient characteristics.	<ol style="list-style-type: none"> 1. Lists facts from the case 2. Groups facts from the case logically 3. Provides reasoning for the groupings of facts 4. Identifies when facts are not relevant to the case
O2. Generate hypotheses or learning questions.	<ol style="list-style-type: none"> 5. Offers hypotheses specific to the case 6. Offers learning questions specific to the case 7. States evidence/reasoning for a given hypothesis or learning question specific to the case 8. Identifies additional data needed to discriminate among hypotheses
O3. Discuss clinical and scientific principles of the disease state.	<ol style="list-style-type: none"> 9. Contrasts epidemiology, pathophysiology or diagnostics of the disease state with normal physiology 10. Explains the pathophysiology of the disease state at the cellular and/or mechanistic level 11. Shares signs and symptoms of the disease and explains how they relate to disease pathophysiology 12. Differentiates between potential causes for signs and symptoms specific to the case 13. Describes an aspect of the clinical course or prognosis of the disease
O4. Discuss clinical and scientific principles of the pharmacotherapies.	<ol style="list-style-type: none"> 14. Lists pharmacologic and/or non-pharmacologic therapy or therapies used in the treatment of the disease <p>Shares in-depth knowledge and is able to apply the standard drug parameters for disease-specific pharmacotherapy (check all that apply below):</p> <ol style="list-style-type: none"> 15. Drug Class/Mechanism of action at the cellular level 16. Usual dosing and/or administration (initial, titration, maximum dose, frequency, all routes of administration, special populations - pregnancy, renal disease, hepatic disease, etc.) 17. Pharmacokinetic and/or pharmacodynamic parameters 18. Adverse effects and/or Drug Interactions 19. Patient education 20. Drug cost and available pharmaco-economic data 21. Contrasts properties within drug classes
O5. Identify Drug-Related Problems (DRPs) through a systematic evaluation process.	<ol style="list-style-type: none"> 22. States/shares and prioritizes DRP based on urgency for resolution of DRP and provides clinical reasoning for level of priority 23. Shares with the group pertinent subjective and/or objective information related to DRP 24. Shares and justifies the presence of a DRP using appropriate literature-based evidence
O6. Formulate a patient-specific, literature-supported plan to resolve DRPs.	<ol style="list-style-type: none"> 25. Shares a plan for resolution of the DRP that includes specific drug, dose, route, frequency, duration and patient education 26. Shares appropriate goals of therapy 27. Shares monitoring parameters for efficacy and toxicity, including frequency of monitoring 28. Shares appropriate literature-based evidence to directly justify proposed plan
O7. Employ an effective strategy to identify relevant scientific literature and resources.	<ol style="list-style-type: none"> 29. Shares and defends literature search terms and/or strategy appropriate to the case by describing databases used, filters or MESH terms, Boolean search logic, etc. 30. Elaborates on the nature of the source used (primary, tertiary, guideline, editorial, etc.)
O8. Evaluate the scientific literature to formulate a solution to the problem.	<p>Appraises the relevance of supporting literature by describing (check all that apply below):</p> <ol style="list-style-type: none"> 31. Identifies current and reputable resources 32. Objectives 33. Study design 34. Population studied

(Continued)

(Continued)

Course Objective and Number (O#) Checklist Item Number

- 35. Outcomes/results
 - 36. Limitations/sources of bias
 - 37. Evaluates external validity of primary literature in relation to the patient case
 - *Overall professionalism performance for each session (Each item must be performed at each session to receive a check mark)
 - 38. Always treats others in the group with dignity and respect
 - 39. Always arrives on time to sessions
 - 40. Always stays the entire session
 - 41. Always demonstrates seriousness of purpose/responsibility for preparedness
 - 42. Always engages others in discussions
 - 43. Always listens attentively to others' opinions
 - 44. Always responds appropriately to others' opinions
 - 45. Always respectfully offers opinions
 - 46. Always provides unprompted information to the group
 - *Global Assessment of the Student's overall performance within a small group over three sessions
 - 47. What overall grade on a scale of 0-100% would you give to this student?
-