
A mobile app-based peer assessment for problem-based learning groups

Received: 7 July 2026

Accepted: 20 August 2026

Published online: 24 August 2026

Cite this article as: Hirman M.M., Bolhassan N.A., Jali N. *et al.* A mobile app-based peer assessment for problem-based learning groups. *BMC Med Educ* (2026). <https://doi.org/10.1186/s12909-026-10244-4>

Melissa Mohammad Hirman, Noor Alamshah Bolhassan, Nurfaiza Jali & William K. Lim

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

A mobile app-based peer assessment for problem-based learning groups**Melissa Mohammad Hirman¹, Noor Alamshah Bolhassan², Nurfaeza Jali², William K. Lim^{1*}**

1. Faculty of Medicine and Health Sciences

Universiti Malaysia Sarawak

94300 Kota Samarahan, Sarawak, Malaysia

2. Faculty of Computer Science and Information Technology

Universiti Malaysia Sarawak

94300 Kota Samarahan, Sarawak, Malaysia

*Corresponding author

William K. Lim

Email: kslim@unimas.my

Abstract**Background:**

Problem-based learning (PBL) has been considered the most influential innovation in medical education for the past half century. This small-group educational approach was first used in medical schools but has since spread to the health sciences and beyond. Self, peer and facilitator assessment have been used for authentic assessment of professional behavior of PBL students. Peer evaluation is a professional skill found to improve student performance. However, peer assessment can be hindered by difficulties such as the lack of anonymity and the problem of friendship marking. Digital tools are increasingly used in education for their accessibility and speed but there is no report on the use of mobile technology in PBL peer assessment.

Methods:

We developed PeerA, an app embedded with rules to limit bias from awarding high scores due to interpersonal relationships. It is designed for confidential weekly grading of peers, with formative feedback in the final week of grading. A trial was done with the Year One Medicine class of 152 students in one 6-week module, followed by an online survey.

Results:

With a response rate of 85%, the results showed that more than two thirds of the students found the app easy to use and agreed to its implementation. Over 70% felt this peer assessment improved their own performance while they benefited from formative feedback provided by their peers. A large majority (94.3%) agreed to the current format of weekly app-based grading followed by feedback in the final week. Furthermore, 65.3% suggested that peer assessment should contribute at least 50% of PBL grades.

Conclusions:

PeerA, a mobile app for quantitative and qualitative peer assessment, provides a feasible and student-endorsed approach for implementing anonymous peer assessment in PBL. Students perceived that its structured rules supported a fairer assessment while encouraging reflective learning. Our findings provide practical guidance for designing digital tools in peer assessment systems for PBL groups.

Keywords: Peer assessment, Problem-based learning, Mobile app, Friendship marking, PBL assessment

Background

Problem-based learning (PBL) was introduced at McMaster University medical school in 1969 as a student-centred, group-based active learning approach that nurtured critical thinking, communication and collaboration. PBL is now considered not only the most influential innovation in medical education of the past fifty years [1] but even as the most innovative instructional method conceived in the history of education [2]. In a scoping review of 124 medical education studies published from 1990 to 2020, PBL was found to be equal or superior to other approaches for knowledge retention and academic performance, and superior at improving social or communication skills, problem-solving and self-learning skills [3].

The ‘McMaster philosophy’ combined three educational components into a unified approach: problem-based, tutored small group and self-directed learning [4]. It was introduced not as a method to add to other methods but as a general educational strategy where learning is student-centred, active, integrated, problem-based, collaborative and tutored [5]. Learning begins when a complex problem is given to a group of commonly, 8-10 students [1]. Students, guided by a facilitator (also called a tutor or teacher in other publications), use their existing knowledge to generate hypotheses and decide what they need to know to fully understand or resolve the problem. They acquire information actively and reconvene to synthesise the knowledge gained. In the process they develop skills such as self-directed learning, problem solving and the so-called 4Cs [6]- creativity, critical thinking, communication and collaboration.

PBL has been described as an aligned teaching system- where the learning outcome is elicited in the teaching method and embedded in the assessment tasks [7]. However, assessment in PBL will be misaligned if it is mainly summative tests of declarative knowledge as these promote rote learning rather than thinking or problem-solving skills [8]. Unlike traditional summative tests for assessing knowledge at the end of each unit of learning, complex abilities cannot be measured at one point in time, but longitudinally and in authentic settings [9]. In the McMaster PBL curriculum, the tutorial-based assessment involving self, peer and facilitator assessment has continued on since the beginning [10]. These have been described as alternative assessments where students take responsibility for their own assessment whereby they reflect, collaborate and dialogue with the teacher. In their review, Dochy, Segers and Sluijsmans stated that self-, peer and co-assessments are particularly relevant to PBL, which is a student-centred setting with repeated small group discussions [11]. Facilitators' feedback alone is not sufficient- the students have to evaluate themselves and their peers [12]. Comparison of self, peer and facilitator assessment in medical PBL have reported a correlation between peer and facilitator scores, but self-ratings tended to be lower or not correlated with facilitator ratings [13]. It is thought that peers give a more accurate picture of a student's PBL performance than the students themselves [14]. As to the benefit of self- assessment, a meta-analysis found self-assessment enhanced academic performance with larger effect sizes if accompanied by feedback from teachers and peers [15]. In facilitator assessment, one facilitator gives a grade to each student whereas in peer assessment each student is evaluated by all the other peers. The use of multiple raters increases the reliability and precision of assessing professional behavior [16]. It is thought that the validity of peer assessment can be demonstrated by how well it compares with assessment by a facilitator [13, 14]. However, studies from medical PBL have found poor-moderate agreement among facilitators scoring the same students [17]. Furthermore, scores differ depending on whether the facilitators were among the lenient or the stringent ones [18]. Some may prefer to give scores close to an average while others allow a wide distribution of scores [11]. In addition to this variability, there is concern that a PBL facilitator cannot possibly play the role of an assessor-evaluator and a partner-mentor at the same time [19]. Peer assessment is considered a generic employability skill that all undergraduate students should develop [20]. In collaborative group settings, where students work together towards a shared learning goal, peer assessment was perceived to benefit student learning more than self-assessment [21]. Peers are in a better position to accurately assess each other's competencies because of the length of time they spend together as compared to the faculty. Much has been published about the positive effects of peer assessment on teamwork and professionalism [22]. Other

benefits include metacognitive skills, critical reasoning, reflection and self-awareness [23]. Nevertheless, there are studies from undergraduate medicine reporting mixed results with peer assessment [23-25]. A recurring obstacle and objection in peer assessment studies is the problem of friendship marking which refers to the tendency of students to give favourable ratings to their friends or close peers. [21, 25, 26]. This can lead to biased assessment with inflated scores arising from a desire to avoid ruining relationships or jeopardising friends. Anonymous peer assessment is one way to avoid such peer pressure [27]. Since anonymity can be compromised by factors such as handwriting, an electronic delivery format is a potential solution [28].

As digital technology permeates all aspects of modern life, it enables learning and assessment to take place without the constraints of space and time. Over the past 20 years, digital tools have become increasingly used in education, enabling digital alternative assessment strategies including technology-supported peer assessment [29]. The advantages include anonymity, speed and efficiency despite large classes, and automatic grade calculation for both quantitative and qualitative feedback [30]. In particular, with mobile technologies students can access digital resources and interact electronically via handheld devices and wireless networks [31]. Mobile technology-supported peer assessment has been reported [32-34] but thus far not in the context of PBL assessment.

At the Universiti Malaysia Sarawak (UNIMAS) medical school, PBL is implemented in the first two years where there are 9 modules of 6-9 weeks each. Each group receives a PBL trigger in the first tutorial where they decide what else they needed to know to fully understand or resolve the problem. They reconvene in the second tutorial to elaborate on their new knowledge and apply it to the problem. Courses are integrated and assessment is based on a balance between content knowledge (modular quiz and examination) and PBL process [35] as assessed by the group facilitator. In the past, this was solely performed by the facilitator. Grades for PBL contribute 5-10% to the grade for each module. Each facilitator grades the students in their group on their performance and participation using a standardised assessment rubric described in Methods. However, there were variations in grading patterns among facilitators. To improve PBL assessment, we aimed to introduce peer assessment alongside facilitator assessment. With a cohort size of 152 students, we needed an electronic delivery format for anonymity and efficiency. Furthermore, we wanted to mitigate the problem of friendship marking. Hence, we developed a mobile application or app for peer assessment (PeerA). After a trial of this app with the students, we surveyed their acceptance of implementing peer assessment alongside facilitator assessment. To our knowledge, this is the first report on the use of mobile technology to implement peer assessment in PBL groups. Therefore, this study aimed to evaluate the

acceptability, feasibility, and perceived educational benefits of our mobile app-based peer assessment system (PeerA) among first-year medical students in a PBL curriculum. Specifically, we sought to answer the following research questions: (1) To what extent do students accept the use of a mobile app for anonymous peer assessment? (2) Does the app-based system mitigate concerns especially friendship marking? (3) What are students' preferences regarding the frequency, format, and weightage of peer assessment relative to facilitator assessment?

Methods

The study was conducted on all 152 first-year medical students during their fifth module which is a 6-week module in June-July 2024. To improve PBL assessment and teach peer evaluation, peer assessment was introduced alongside facilitator assessment. A trial of peer assessment (without the app) was conducted on the student cohort immediately prior to this one and the rules for this trial came partially from the feedback of that cohort. The students were briefed on peer assessment as a transversal skill at the start of their course and practiced using the app over their first 4 PBL-based modules. They were then enrolled into the study and graded their peers weekly for 5 weeks while the 6th week was for formative feedback. Subsequently they were surveyed for their responses.

Development of the mobile app for peer assessment

The PeerA app was a collaborative effort between the Faculty of Computer Science and Information Technology, the Faculty of Medicine and Health Sciences and the Tun Abang Haji Openg Digital Centre of UNIMAS. The collaboration bridged educational objectives with technical capabilities, ensuring that the resulting system would be pedagogically sound, user-friendly, and operationally efficient. It was developed using 'REKA', the App Creator platform conceived at UNIMAS (<https://ia.unimas.my/>).

Access control

Each user is assigned to one of three predefined roles: student, facilitator, or coordinator. Students can log in to submit weekly peer assessments and qualitative feedback, and view only their own cumulative results. They do not have access to other users' data and cannot modify previously submitted entries once the grading window closes. Facilitators are allowed to view aggregated results within their assigned PBL groups. They can monitor student participation, authorise absences, and download summary reports for their respective sessions. Above the facilitators, coordinators oversee the overall peer assessment activity where they can configure academic years, manage blocks and group structures, monitor overall assessment statistics, and export data for institutional analysis.

Functions of the peer assessment app

The PeerA app is an easy-to-use software where students can anonymously grade each student in the group. The grading criteria are provided within the app which are similar to those used in facilitator assessment. As depicted in Figure 1, the grading criteria are grouped under 4 domains- participation and contribution, group skills and teamwork, communication and presentation skills, and attendance and punctuality. The app automatically calculates the average weekly marks as well as the final module mark which is the average of the 5 weekly marks. Students can only view their own average grade and not the marks awarded by each peer. To ensure participation, any student who did not grade for a week will receive a zero grade for that week. This does not affect the marks of the other peers because the system auto-corrects for the total number of students who graded that week when calculating the average mark. To avert overly high grades related to friendship marking, students were only allowed to give the full mark of 10 to a maximum of two students, while all other marks (1-9) could only be given a maximum of 4 times each. Grading was opened after the week's final PBL session and closed at the start of the subsequent week's PBL session. On the final week of the module, the students do not grade but give qualitative feedback to each peer by filling in 'One thing I appreciate about you' and 'One thing I request from you'.

Peer Assessment

Rubrics

1. PBL Participation & Contribution

- Displayed keen interest to learn
- Participated actively in all sessions
- Played good leadership role when required
- Always contributed useful, logical, complete & well organised information

0 1 2 3 4

2. Group Skills and Teamwork

- Polite & supportive to others
- Encouraged other members to participate
- Listened actively & paid attention to input provided by others
- Keen to share knowledge with others but not dominating the session

0 1 2 3 4

3. Communication & Presentation Skills

- Gives logical and clear explanations easily understood by others
- Used & utilised aids to enhance clarity of content and retention of information

0 1 2

4. Punctuality

- Arrived on time unless with good reason

0 1

Cancel Save

Fig 1. Screenshot of the criteria for student peer assessment

Student survey and data analysis

On completion of the module, students were surveyed anonymously with a mixed-format questionnaire using a Google Form. All the students consented to participate in the study and to its publication. The first section of the survey relates to acceptance of peer assessment, using a 5-point Likert scale. The second section included 8

questions on the various processes of peer assessment to which the respondent had to select one out of several options. Open-ended questions were included to obtain the students' perceptions and recommendations. The responses were analyzed using Microsoft Excel® (2013) and the response frequencies converted to percentages. Open-ended responses were grouped together according to similar themes.

Results

Following their 6-week trial, the students were surveyed for their response to the PeerA-based peer assessment. One hundred and twenty-seven students responded to the survey giving a response rate of 84.7%.

Acceptance of peer assessment and use of PeerA app

To gauge their acceptance of the PeerA-based peer assessment, students were asked to choose from 5 options ranging from strongly agree to strongly disagree, if this should be implemented in all the modules. The majority strongly agreed or agreed that it should be implemented (66.2%), with 28.4% strongly agreeing (Figure 2). Most of the students rated strongly agree or agree that they improved in their ability to grade peers (69.3%). More than two-thirds of the students (70.8%) found the PeerA app useful and easy to use. Vast majorities stated they had improved their own performance by grading their peers (74.1%) and by reading the qualitative feedback from their peers (74.8%). More than half of the students (59.0%) agreed that the app reduced friendship marking.

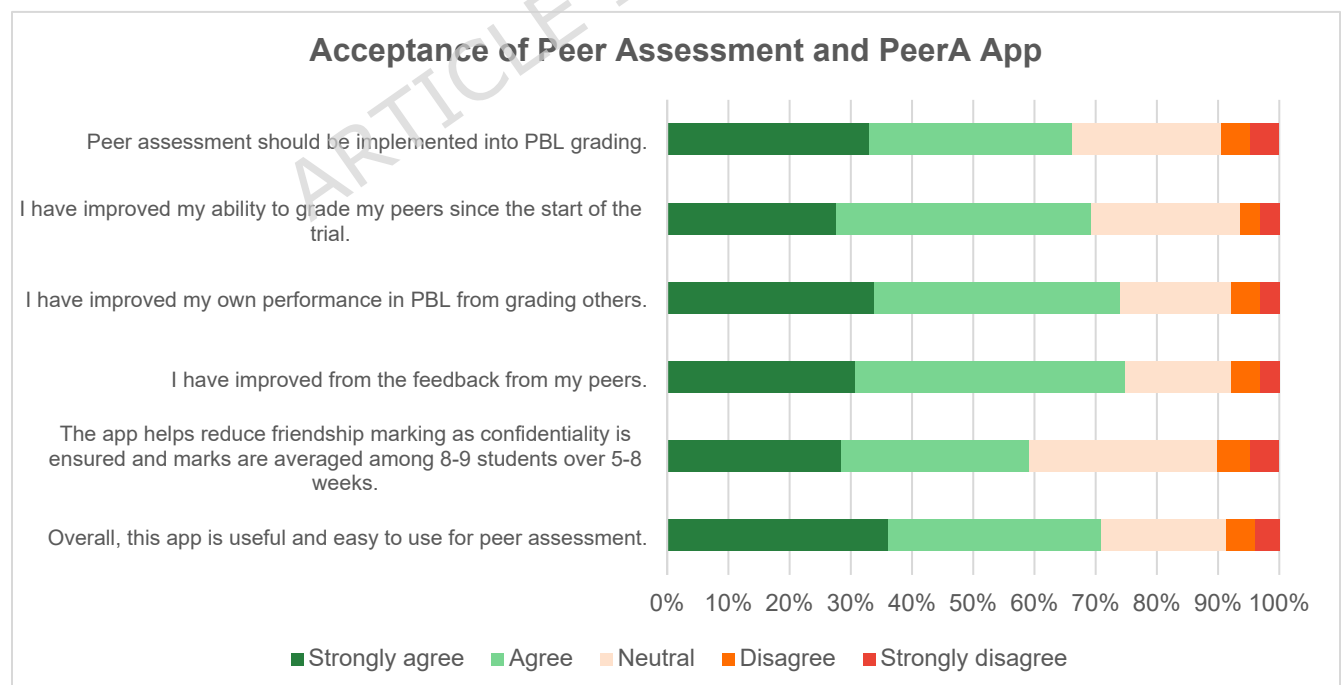


Fig. 2 Student's acceptance on peer assessment and PeerA App in PBL

Acceptance of peer grading criteria and policies

As shown in Table 1, the students unanimously agreed to retain the grading criteria. When asked if they agree to the limit on awarding full marks, almost half (41.7%) replied in the affirmative. However, a higher percentage wanted to raise this limit to 3 (26.8%) and 4 (22.8%) respectively. We had set the rule that every mark (aside from 10) can only be given to 4 students at most. Almost half the students (48.0%) agreed to this limit although a smaller percentage (29.1%) wanted to raise it to 5. The vast majority (94.3%) voted to maintain the current format of weekly grading followed by qualitative feedback in the final week.

Table 1 Student's views on the grading processes in peer assessment

Survey Question	Frequency (%)
<i>Do you agree with the current criteria for peer assessment?</i>	
Yes	127 (100.0)
No	0 (0.0)
<i>How many people do you think should be able to receive the full marks of 10?</i>	
1	3 (2.4)
2 (current)	53 (41.7)
3	34 (26.8)
4	29 (22.8)
Others	8 (6.3)
<i>If you choose Other, type the number of people that you think can receive the full mark of 10.</i>	
0	0 (0.0)
5	5 (62.5)
all	1 (12.5)
No limit	2 (25.0)
<i>Aside from 10 marks, what is the maximum number of times one particular mark can be given?</i>	
1	0 (0.0)
2	1 (0.8)
3	4 (3.1)
4 (current)	61 (48.0)
5	37 (29.1)
6	13 (10.2)
7	0 (0.0)
8	11 (8.7)
<i>In a 6-week Block, how many weeks should be used for grading (currently it is 5 weeks grading and 1 week feedback)?</i>	
1	0 (0.0)
2	2 (1.6)
3	5 (3.9)
4	3 (2.4)
5 (current)	117 (92.1)

Regarding how many weeks should be used for grading in a 6-week block, almost all students (92.1%) chose the current 5 weeks grading. When asked the reason for this choice, students said it is a fairer approach, enables continuous improvement and is easier to remember (Table 2).

Table 2 Reasons for choosing a 5-week grading and 1-week feedback

Theme	Sample Comments
Accurate and fair assessment	<i>Everyone performance is different for each week (...). There may be times when the students are unable to properly present their discussion points (...) with a higher number of grading weeks, everyone will have to learn. (...) have sufficient time to observe and assess the student's performance.</i>
Allows continuous improvement	<i>(...) make students more focus in preparation for the PBL. (...) it allows the peers to notice the mistake and improving for the next week (...)</i>
Convenience	<i>(...) students will start their study week and sometimes, we tend to forget (...) when they forget to comment on week 6, it will not affect the overall marks.</i>

Decision on facilitator-peer assessment weightage

This peer assessment is the first time students have an input into PBL grades, raising the question of what proportion of the grades should come from the students. The students were asked if peer assessment should contribute 0% to PBL grade (the status quo) or 100% (giving no role for the facilitator) or a joint contribution where students contribute 25%, 50% or 75% of the total grade. The results in Figure 3 show the largest single group of 39.4% voted for half the PBL assessment to come from peer evaluation, with only 15% wanting students to have a higher weightage. Although only 2.9% wanted the status quo of full facilitator grading, a very significant 45.7% wanted more marks to come from facilitators than students. This exceeds the 39.4% who wanted equal weightage between facilitators and students and argues for a final weightage that slightly favors facilitators over students.

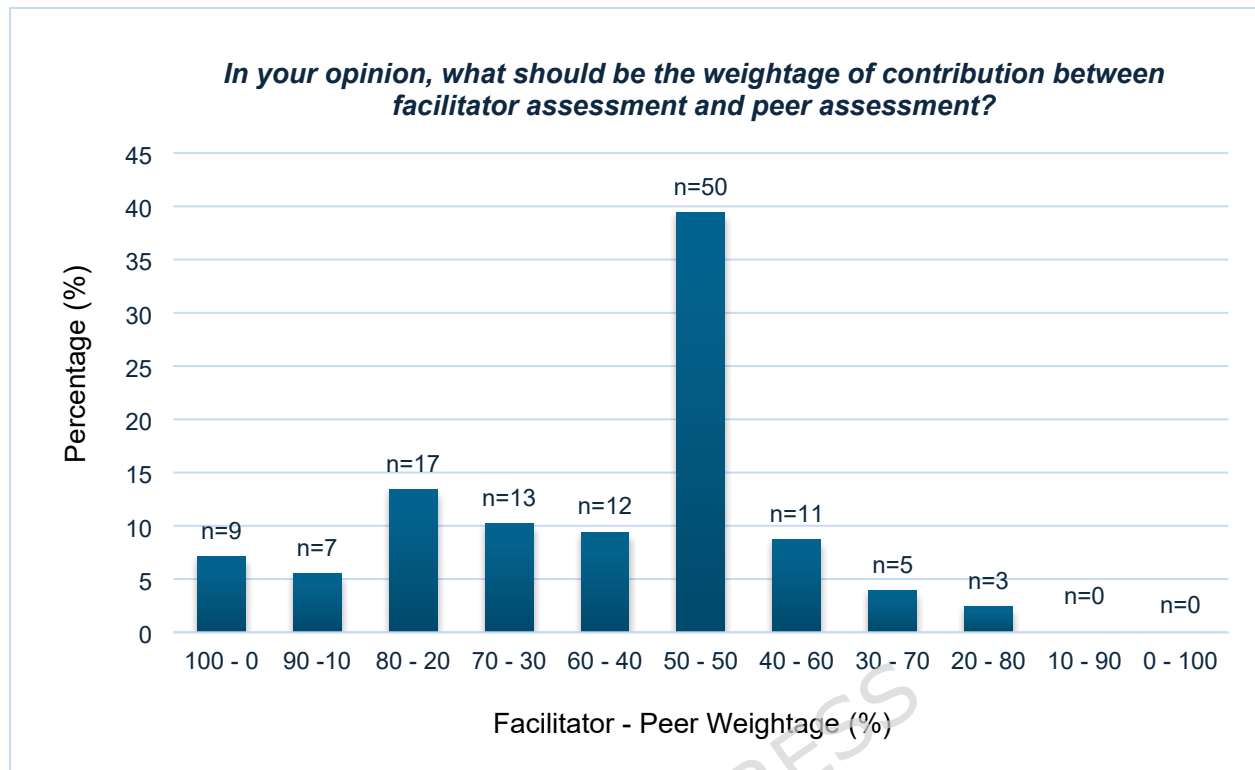


Fig 3 Students' preferences for facilitator versus peer weightage in PBL assessment

Views on quantitative and qualitative feedback

We integrated quantitative (weekly) with qualitative (final week only) feedback so the students can obtain constructive suggestions for improvement in the next module. The qualitative feedback was structured by way of 2 statements: (i) One thing I appreciate about you (ii) One thing I request from you. These 2 statements were agreed upon by vast majorities of 83.2% and 92.1% respectively (Figure 4).

Interestingly, a large majority (81.1%) wanted qualitative feedback to be available not just in the final week but every week.

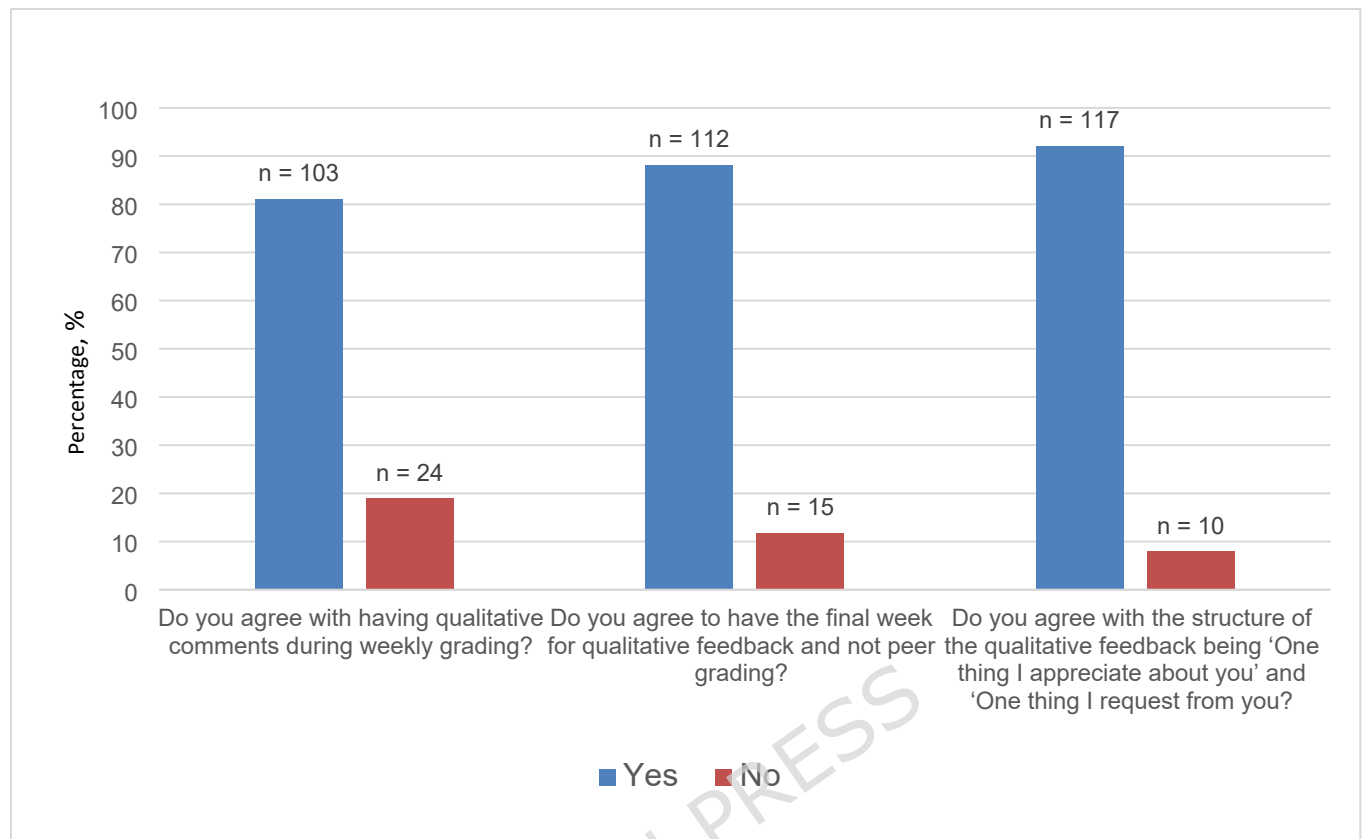


Fig 4 Students' preferences on quantitative and qualitative feedback format

Additionally, the students were asked for their views on quantitative and qualitative feedback, as listed in Table 3.

Reasons for suggesting peer grading in every single week of the module include better accuracy, fairness and continuity. Regarding the qualitative feedback format, some thought reading 'One thing I request from you' from peers could produce anxiety and stress in PBL sessions. Others suggested merging the 2 statements for simplicity.

Currently the qualitative feedback in the final week of the module is intended as a stimulus for the group to conduct a debriefing for improving themselves in the next module. One student suggested incorporating the writing of qualitative comments together with the weekly grading.

Table 3 Student's open comments on quantitative and qualitative feedback

Type of feedback	Theme	Sample Comments
Weekly quantitative grading excluding final week	Accuracy and fairness	(...) to provide a more average result for the block with more grading mark provided. (...) there's a possibility of someone getting really low marks in the first few weeks but performs well in the final week. (...) because it's not fair to chairperson and also the one who are actively involve is not being evaluated (...)
	Continuity	Every PBL will be important in our learning (...). (...) feels like disrupting the continuity of the assessment.

Qualitative feedback on final week	Preference to positive feedback	(...) <i>"One thing I request from you" might cause someone to overthink on how he/she should improve (...) might cause someone unable to share information efficiently (...) because he/she is afraid of the what the peers thought of his/her performance.</i> (...) <i>positive words from group members will motivate us to do better for upcoming discussions.</i>
	Simplicity	(...) <i>this 2 can actually combine as one.</i> <i>Only 1 space for writing should be enough.</i>

Views on training and other open-ended feedback

The training of students for peer assessment is said to increase the quality of assessment and positive attitudes towards it [29]. When asked if they need training in peer assessment, a large majority (n = 107, 84.3%) replied in the negative. Of the remaining 20 respondents who requested training, some of the suggested topics included *'How to grade fairly'*, *'How to differentiate between people who perform equally good'* and *'What are the basic criteria of a good student'*. Under Open Comments, students gave a variety of responses including wariness about insincere graders (e.g. *'My classmates are not genuine in writing comments....'* *Certain people might sabotage our marks.'*) and affirmation of learning more from PBL because of the peer assessment (e.g. *'I like how I actually learned more in PBL'*, *'Throughout the first year, I'm satisfied with the improvement.'*).

Discussion

We introduced peer assessment to enable students to take greater ownership of their assessment and develop peer evaluation skills. However, peer assessment may be hindered by poor evaluation skills, friendship marking leading to unfairness and inaccuracy, and the lack of anonymity or efficiency in its implementation [22, 23, 36, 37]. We developed a mobile app (PeerA) to implement PBL peer assessment in way that addressed these challenges.

Anonymous assessors are more likely to give honest feedback [38]. With the app, there is no increase in administrative workload as the marks are calculated automatically and accurately. About three quarters said peer assessment improved their own PBL performance, consistent with a study that showed those who assessed peers gave better performance in their small groups [39].

Student acceptance

After a 6-week trial of this PeerA-based peer assessment, more than two thirds of the students found the app easy to use and agreed for it to be implemented. A similar proportion of the students had found the app to be user-friendly and had rising confidence in grading over time. Another factor could be the fact that these first-year students had

only a short exposure to facilitator grading. An earlier smaller trial using Google Form on a second-year cohort received some negative responses akin to a published report that peer assessment rendered the friendly and harmonious atmosphere of the PBL group to one that was threatening, competitive, with a fear of being judged [25]. In this cohort, approximately 10% voted their disagreement with peer assessment. From the qualitative feedback, some students trusted the facilitator more than their peers. A future survey will be able to show whether confidence in peers will increase after several more cohorts become used to peer assessment.

It has been reported that student ownership of the assessment criteria is fundamentally important to its success [26]. Meer and Chapman found for students to fully understand the marking criteria, they need to actively participate in its design [40]. We implemented peer assessment on the cohort using the rules that the majority in that cohort had approved.

Support and training

As this change would be new to many, it is necessary to train students in assessing competencies. Scaffolding strategies to support students include guidance on the assessment rubrics [12]. The PeerA app provides scaffolding for grading as it displays every rubric for the students to assign a mark to each rubric. Most of the students felt they did not need further training, suggesting they had the confidence to grade their peers using the app. Those who felt they needed further training stated areas such as knowing how to grade fairly and the criteria of a good student. These can be part of the training of future cohorts.

Frequency of assessment

One of the strengths of our peer assessment is that the PBL group members stay together for 2 years. Hence the students can evaluate each other more accurately due to the long period of observation and interaction. Grading was done every week of the module up to the final week where it was replaced with formative feedback. A review of technology-supported peer assessment found the majority of studies implemented only one round of assessment [42]. The assessment of skills is more reliable when done over multiple times as not all skills may be demonstrated in any one session. Students may require time to adjust to one another to collaborate and perform their tasks effectively. Further, the students' ability to appraise professional behavior will rise over time [16]. The necessity for multiple rounds of assessment was supported by a study that showed the initial feedback was more cognitive, followed by metacognitive evaluation in the second round and ending with affective responses in the third round [43]. The vast majority of our students (92.1%) agreed with weekly grading. Their comments indicated weekly

assessment was more accurate and fairer (e.g. *'Everyone's performance is different for each week, so 5 weeks is good enough to take the average'*). Students also opined time was needed for improvement, with comments such as *'... it allows peers to notice the mistake and improving for the next week...'* and *'Help us to keep improving'*. Some even wanted to grade during the final week to preserve the continuity of the assessment (e.g. *'...there's a possibility of someone getting really low marks in the first few weeks but performs well in the final week'*). However, our system of giving a final average mark should compensate for any lower performance in a particular week.

Friendship marking

Social relations among students can render peer assessment invalid due to the influence of enmity or friendship. Brown and Knight listed 4 rating errors [44] of which the first is friendship marking. As described in Background, this is a cooperative arrangement of giving inflated marks to maintain certain friendships and interpersonal relationships among peers [22, 25, 45]. It can also manifest as having many students receive similar high marks despite a wide variance in their actual performance [46]. This causes unfair and invalid grading and can arise out of a fear of peers receiving bad grades as well as a desire for social acceptance. Our app mitigates this in several ways. Firstly, the weekly grade is the average of the grades given by every peer while the final grade is the average of these weekly grades. This helps to balance higher grades given by certain peers, or in certain weeks. Secondly, the grading is anonymous – the student is shown the average mark for the week but does not know which peer gave what mark. Thirdly, we modified a method first reported in team based learning to prevent indiscriminate high marks for every student [36]. The highest mark can only be awarded twice and other marks not more than four times. Any deviation from this will halt the grading and activate a reminder message from the app. This ruling was supported by almost 42% of the students (the biggest single grouping). However, those who wanted to give full marks more than twice totaled up to almost 56%. Some gave feedback such as *'Please give main criteria to determine good peer'* and *'I guess should have put 4 people getting 10'*. This may indicate a need for more extensive training on how to perform peer evaluation. Hence, we demonstrate friendship marking can be mitigated by anonymity, averaging of the marks and applying rules that prevent indiscriminate awarding of high grades.

Qualitative feedback

While we used numeric scales for weekly summative numerical grading, the final week of the module was dedicated to structured formative feedback. The first part was positive feedback (one thing I appreciate about you) and the second was constructive advice for improvement (one thing I request from you). College students were reported to

be able to give valid feedback with the proportion of valid or useful feedback exceeding 80% [47]. It was also reported that while feedback was more general, feedback from students was more specific, and thus the two complemented each other. A review of technology-supported peer assessment studies found that a majority of 61% consisted of both quantitative and qualitative feedback [42]. A vast majority of over 90% of our students agreed with both the current peer assessment format as well as our structured feedback format. Interestingly, a big majority of 81.1% wanted the opportunity to give qualitative feedback every week. This could indicate a desire to provide justification for their scores and receive more input for faster improvement. There were also some critical comments on providing narrative feedback (e.g. '*...it took a lot to think...*', '*I think it easier to not use long wordy feedback method*'). Correspondingly, 15.7% requested further training, showing the giving of constructive feedback is not intuitive but a skill to be developed.

Deciding between peer and facilitator assessment

Students spend more time with each other than with faculty, so peer assessment is based on more careful observation. A meta-analysis on the impact of peer assessment found that it improves academic performance compared to facilitator assessment, despite the greater qualifications and pedagogical experience of the teacher [38]. Facilitators have been reported to regard students as too immature to be assessors [45]. Two of our students expressed doubt about their peers with the comments that '*Certain people might sabotage our marks*' and '*My classmates are not genuine when writing comments about performance, so I don't know how to improve*'. Yet, PBL assessment by facilitators have been reported to have low inter-facilitator reliability [17], with high variability between the stringent and lenient facilitators [18]. When given the choice between facilitator assessment and peer assessment, and the weightage between the two, none of our students wanted solely peer assessment, while 7% chose full facilitator assessment. However, the biggest single bloc of students (39%) chose a 50:50 share of the scores between facilitators and peers. Overall, more preferred a higher weightage for facilitators (45.6%) versus more marks for peers (15%). In recognition of this trend, we proposed a 70:30 share with higher weightage for facilitators. In principle, a heavier weightage for peers would better fit PBL as a student-centred learning approach. However, we expect the ability of students to evaluate peers to improve over time [11]. It may be that our future student cohorts will support a shift towards a heavier weightage for peer assessment. Nevertheless, despite the general positive views on peer assessment and the use of PeerA app, a small percentage of students had negative views and raised concerns on fairness and genuineness of peer feedback. This suggests trust remains an important

issue for some students. As requested by the students themselves, further training and experience with peer assessment may help build confidence in the process over time.

Limitations and future directions

This work is largely descriptive and serves as a basis to implement summative and formative PBL peer assessment. Our findings may be more relevant to small group-based pedagogies where students are able to evaluate peers repeatedly over several weeks. However, the benefits of using a mobile app for peer assessment should have wide relevance. Although the app was set up to mitigate friendship marking, collusive marking is still a possibility, where students use a grading formula to ensure similar grades for every peer. To address this, it is necessary for frequent monitoring of the grades, while penalties applied on rule-breakers serve as a warning to the rest of the class. Future work should investigate the correlation between facilitator grading and peer assessment. A positive correlation will indicate that peer assessment is as reliable as facilitator grading in the evaluation of PBL performance. This will enable facilitators to have more confidence in grading by students. Self-assessment could also be implemented to evaluate its reliability within our student cohort. Moving forward it is clear students could benefit from more training, particularly in how to give specific and constructive feedback, using good examples as a guide. As students become more skilled and experienced with peer assessment, it is anticipated they will be more confident in student grading and support a greater weightage of marks towards peer assessment.

Conclusion

Peer assessment is recognised for its benefits in improving students' academic performance but is often challenged with problems of accuracy and bias. Our students generally agreed that the PeerA-based peer assessment benefited their learning and the processes implemented supported fair and accurate grading. The mobile app was found easy to use, supported anonymity and served as a useful tool for conducting online peer assessment in PBL settings.

Overall, our findings support the use of digital tools to promote fairer and more accurate grading in peer assessment, both in general and for PBL in particular.

List of Abbreviations

PBL: Problem-based learning

UNIMAS: Universiti Malaysia Sarawak

Declarations

Ethics approval and consent to participate

The ethical approval for this study was granted by the Medical Research Ethics Committee of Universiti Malaysia Sarawak (FME/24/135). Informed consent was obtained from all individual participants included in the study including publication of data.

Consent for publication

Not applicable

Availability of data and materials

The data generated or analysed in this study are included in this article. Other information is available from the corresponding author on reasonable request.

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Funding

No funding was involved in the conduct of this study.

Authors' contributions

MH designed and administered the questionnaire, curated and analysed the data, and contributed to the writing and revision of the manuscript. NA led the development of the PeerA application and contributed to the study methodology. NJ contributed to the methodological design and conducted student training on the use of the PeerA application. WL conceptualised and supervised the study and contributed to the writing and critical revision of the manuscript. All authors read and approved of the final manuscript.

Acknowledgements

We thank Ms. Sharifah Norizan Wan Zain for technical input in the project and acknowledge Mr. Mohd Razif Baital@Latif who developed the App Creator platform used to build the PeerA application. We thank Universiti Malaysia Sarawak for approval and administrative support of the study.

References

- [1] Wood DF. Problem based learning. *BMJ*. 2008;336:971.
- [2] Woei Hung DHJ, Rude Liu. Problem-Based Learning. In: David Jonassen MJS, Marcy Driscoll, M. David Merrill, Jeroen van Merriënboer, Marcy P. Driscoll, editor. *Handbook of Research on Educational Communications and Technology*. 3rd Edition ed. New York: Routledge; 2008. p. 1296.
- [3] Trullas JC, Blay C, Sarri E, Pujol R. Effectiveness of problem-based learning methodology in undergraduate medical education: a scoping review. *BMC Med Educ*. 2022;22:104.

- [4] Neufeld VR, Barrows HS. The "McMaster Philosophy": an approach to medical education. *J Med Educ.* 1974;49:1040-50.
- [5] Walton HJ, Matthews MB. Essentials of problem-based learning. *Med Educ.* 1989;23:542-58.
- [6] Thornhill-Miller B, Camarda A, Mercier M, Burkhardt J-M, Morisseau T, Bourgeois-Bougrine S, et al. Creativity, Critical Thinking, Communication, and Collaboration: Assessment, Certification, and Promotion of 21st Century Skills for the Future of Work and Education. *J. Intell.* 2023;11:54.
- [7] Biggs J. What the Student Does: teaching for enhanced learning. *High. Educ. Res. Dev.* 1999;18:57-75.
- [8] Waters R, McCracken M. Assessment and evaluation in problem-based learning. *Proceedings Frontiers in Education 1997 27th Annual Conference Teaching and Learning in an Era of Change.* 1997;2:689-93 vol.2.
- [9] van der Vleuten CPM, Schuwirth LWT. Assessment in the context of problem-based learning. *Adv Health Sci Educ Theory Pract.* 2019;24:903-14.
- [10] Neville A, Norman G, White R. McMaster at 50: lessons learned from five decades of PBL. *Adv Health Sci Educ Theory Pract.* 2019;24:853-63.
- [11] Sluijsmans DMA, Moerkerke G, van Merriënboer JJG, Dochy FJR. Peer assessment in problem based learning. *Stud. Educ. Eval.* 2001;27:153-73.
- [12] Concina E. The Relationship between Self- and Peer Assessment in Higher Education: A Systematic Review. *Trends High. Educ.* 2022;1:41-55.
- [13] Papinczak T, Young L, Groves M, Haynes M. An analysis of peer, self, and tutor assessment in problem-based learning tutorials. *Med Teach.* 2007;29:e122-32.
- [14] Kamp RJ, Dolmans DH, Van Berkel HJ, Schmidt HG. Can students adequately evaluate the activities of their peers in PBL? *Med Teach.* 2011;33:145-50.
- [15] Yan Z, Wang X, Boud D, Lao H. The effect of self-assessment on academic performance and the role of explicitness: a meta-analysis. *Assess. Eval. High. Educ.* 2023;48:1-15.
- [16] Cottrell S, Diaz S, Cather A, Shumway J. Assessing Medical Student Professionalism: An Analysis of a Peer Assessment. *Med. Educ. Online* 2006;11:4587.
- [17] Khawaji B, Masuadi E, Alraddadi A, Khan MA, Aga SS, Al-Jifree H, et al. Tutor assessment of medical students in problem-based learning sessions. *J Educ Health Promot.* 2024;13:237.
- [18] Sa B, Ezenwaka C, Singh K, Vuma S, Majumder MAA. Tutor assessment of PBL process: does tutor variability affect objectivity and reliability? *BMC Med Educ.* 2019;19:76.
- [19] Servant-Miklos VFC. Fifty Years on: A Retrospective on the World's First Problem-based Learning Programme at McMaster University Medical School. *Health Prof. Educ.* 2019;5:3-12.
- [20] Cassidy S. Developing employability skills: peer assessment in higher education. *ET.* 2006;48:508-17.
- [21] Stančić M. Peer assessment as a learning and self-assessment tool: a look inside the black box. *Assess. Eval. High. Educ.* 2021;46:852-64.
- [22] Lerchenfeldt S, Mi M, Eng M. The utilization of peer feedback during collaborative learning in undergraduate medical education: a systematic review. *BMC Med Educ.* 2019;19:321.
- [23] Curran VR, Fairbridge NA, Deacon D. Peer assessment of professionalism in undergraduate medical education. *BMC Med Educ.* 2020;20:504.

- [24] Roberts C, Jorm C, Gentilcore S, Crossley J. Peer assessment of professional behaviours in problem-based learning groups. *Med Educ.* 2017;51:390-400.
- [25] Papinczak T, Young L, Groves M. Peer assessment in problem-based learning: a qualitative study. *Adv Health Sci Educ Theory Pract.* 2007;12:169-86.
- [26] Pond K, ul-Haq R. Learning to assess students using peer review. *Stud. Educ. Eval.* 1997;23:331-48.
- [27] Vanderhoven E, Raes A, Montrieux H, Rotsaert T, Schellens T. What if pupils can assess their peers anonymously? A quasi-experimental study. *Comput. Educ.* 2015;81:123-32.
- [28] Finn GM, Garner J. Twelve tips for implementing a successful peer assessment. *Med Teach.* 2011;33:443-6.
- [29] Helden GV, Werf VVD, Saunders-Smiths GN, Specht MM. The Use of Digital Peer Assessment in Higher Education—An Umbrella Review of Literature. *Ieee Access.* 2023;11:22948-60.
- [30] Zheng B, He Q, Lei J. Informing factors and outcomes of self-assessment practices in medical education: a systematic review. *Ann Med.* 2024;56:2421441.
- [31] Hwang G-J, Wu P-H. Applications, impacts and trends of mobile technology-enhanced learning: a review of 2008-2012 publications in selected SSCI journals. *Int J Mob Learn Organisation.* 2014;8:83-95.
- [32] Chang C, Lin H-CK. Effects of a mobile-based peer-assessment approach on enhancing language-learners' oral proficiency. *Innov. Educ. Teach. Int.* 2020;57:668-79.
- [33] Hsia L-H, Sung H-Y. Effects of a mobile technology-supported peer assessment approach on students' learning motivation and perceptions in a college flipped dance class. *Int J Mob Learn Organ.* 2020;14:99–113.
- [34] Ramesh JVN, Alazzam MB, Mahdi HS, Sulaiman SK, Sharif SFA, Raj II. Mobile-Based Peer Assessment in English Language Learning: Facilitating Collaborative Feedback in ESL Classrooms. 2025 3rd International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC)2025. p. 672-7.
- [35] Lloyd-Jones G, Margetson D, Bligh JG. Problem-based learning: a coat of many colours. *Med Educ.* 1998;32:492-4.
- [36] Farland MZ, Sicat BL, Franks AS, Pater KS, Medina MS, Persky AM. Best practices for implementing team-based learning in pharmacy education. *Am J Pharm Educ.* 2013;77:177.
- [37] Abraham R, Singaram VS. Self and peer feedback engagement and receptivity among medical students with varied academic performance in the clinical skills laboratory. *BMC Med Educ.* 2024;24:1065.
- [38] Double KS, McGrane JA, Hopfenbeck TN. The Impact of Peer Assessment on Academic Performance: A Meta-analysis of Control Group Studies. *Educ. Psychol. Rev.* 2020;32:481-509.
- [39] Schonrock-Adema J, Heijne-Penninga M, van Duijn MA, Geertsma J, Cohen-Schotanus J. Assessment of professional behaviour in undergraduate medical education: peer assessment enhances performance. *Med Educ.* 2007;41:836-42.
- [40] Meer N, Chapman A. Co-creation of Marking Criteria: Students as Partners in the Assessment Process. *BMHE* 2015:1-15.
- [41] Ibarra-Sáiz MS, Rodríguez-Gómez G, Boud D. Developing student competence through peer assessment: the role of feedback, self-regulation and evaluative judgement. *High. Educ.* 2020;80:137-56.

- [42] Zheng L, Chen N-S, Cui P, Zhang X. A Systematic Review of Technology-Supported Peer Assessment Research: An Activity Theory Approach. *Int. Rev. Res. Open Distrib. Learn.* 2019;20:168-91.
- [43] Tsai C-C, Liang J-C. The development of science activities via on-line peer assessment: the role of scientific epistemological views. *Instr. Sci.* 2009;37:293-310.
- [44] Brown S, Knight P. *Assessing Learners in Higher Education*. 1st edition ed. London: Routledge; 1994.
- [45] Bollela VR, Gabarra MHC, da Costa C, Lima RCP. Students and tutors' social representations of assessment in problem-based learning tutorials supporting change. *BMC Med Educ.* 2009;9:30.
- [46] Heyman JE, Sailors JJ. Peer assessment of class participation: applying peer nomination to overcome rating inflation. *Assess. Eval. High. Educ.* 2011;36:605-18.
- [47] Rollinson P. Using peer feedback in the ESL writing class. *ELT J.* 2005;59:23-30.



Peer Assessment

Rubrics

1. PBL Participation & Contribution

- Displayed keen interest to learn
- Participated actively in all sessions
- Played good leadership role when required
- Always contributed useful, logical, complete & well organised information

0	1	2	3
---	---	---	---

2. Group Skills and Teamwork

- Polite & supportive to others
- Encouraged other members to participate
- Listened actively & paid attention to input provided by others
- Keen to share knowledge with others but not dominating the session

0	1	2	3	4
---	---	---	---	---

3. Communication & Presentation Skills

- Gives logical and clear explanations easily understood by others
- Used & utilised aids to enhance clarity of content and retention of information

0	1	2
---	---	---

2. Group Skills and Teamwork

- Polite & supportive to others
- Encouraged other members to participate
- Listened actively & paid attention to input provided by others
- Keen to share knowledge with others but not dominating the session

0	1	2	3	4
---	---	---	---	---

3. Communication & Presentation Skills

- Gives logical and clear explanations easily understood by others
- Used & utilised aids to enhance clarity of content and retention of information

0	1	2
---	---	---

4. Punctuality

- Arrived on time unless with good reason

0	1
---	---

Cancel

Save