Spring 3-1-2016

Reliability assessment of a peer evaluation instrument in a team-based learning course

Joy Wahawisan

Miguel Salazar

Robin Walters

Fadi M. Alkhateeb

Omar F. Attarabeen
Marshall University, attarabeen@marshall.edu

Recommended Citation
Wahawisan J, Salazar M, Walters R, Alkhateeb FM, Attarabeen O. Reliability assessment of a peer evaluation instrument in a team-based learning course. Pharmacy Practice. 2016 Jan-Mar;14(1):676. ;14(1):676. doi: 10.18549/PharmPract.2016.01.676

Wahawisan J, Salazar M, Walters R, Alkhateeb FM, Attarabeen O. Reliability assessment of a peer evaluation instrument in a team-based learning course. Pharmacy Practice 2016 Jan-Mar;14(1):676. doi: 10.18549/PharmPract.2016.01.676

Original Research

Reliability assessment of a peer evaluation instrument in a team-based learning course

Joy WAHAWISAN, Miguel SALAZAR, Robin WALTERS, Fadi M. ALKHATEEB, Omar ATTARABEEN.

Received (first version): 5-Oct-2015 Accepted: 1-Mar-2016

ABSTRACT[*]

Objective: To evaluate the reliability of a peer evaluation instrument in a longitudinal team-based learning setting.

Methods: Student pharmacists were instructed to evaluate the contributions of their peers. Evaluations were analyzed for the variance of the scores by identifying low, medium, and high scores. Agreement between performance ratings within each group of students was assessed via intra-class correlation coefficient (ICC).

Results: We found little variation in the standard deviation (SD) based on the score means among the high, medium, and low scores within each group. The lack of variation in SD of results between groups suggests that the peer evaluation instrument produces precise results. The ICC showed strong concordance among raters.

Conclusions: Findings suggest that our student peer evaluation instrument provides a reliable method for peer assessment in team-based learning settings.

Keywords: Students, Pharmacy; Education, Pharmacy; Educational Measurement; Curriculum; Cooperative Behavior; United States

INTRODUCTION

Team-based learning is a small-group collaborative process that occurs during class, utilizing individual preparation with a focus on application of learned materials. Within a student team-based learning environment, peer collaboration occurs regularly. A significant portion of a student's individual grade is determined based upon group effort, although it is not realistic that each student will consistently contribute an amount that everyone perceives to be equal. To ameliorate this perceived uneven distribution of effort, peer evaluation instruments are utilized. There are many ways to conduct a peer evaluation process, each with their own pros and cons; however there is limited data on the objective reliability of the instruments, as well as the subjective efficacy of the evaluation process on changes in team dynamics.[1]

Team-based learning (TBL) is a teaching approach that is increasingly employed in higher education. According to a study by Allen *et al*., faculty among approximately one-third of US schools and colleges of pharmacy have carried out team-based learning in stand-alone courses or across the curriculum.[2] The reason for more frequent utilization of TBL is that it employs active learning in order to promote self-motivated scholarship. Furthermore, TBL enhances student communication skills and establishes a team-based environment, which promotes problem-solving capabilities for students.[3] Additionally, TBL is beneficial to course facilitators as students are partially supervised by their peers. This supervision builds up a sense of responsibility towards their individual and team performance. Consequently, learning and scholarship improve among all students regardless to their previous academic performance or grades history.[4-6] When TBL is employed, students perform better on examination questions, indicating their increased mastery of course content.[7-9] Previous research showed that utilizing TBL helps students achieve the same or better knowledge scores as compared to using more traditional learning methods.[10]

At the Rangel College of Pharmacy, team-based learning is utilized in courses focused towards application of knowledge to create patient care plans in an interactive process. If students have questions during their creation of the care plan, they have the opportunity to ask their clinical facilitators for explanation or guidance. Additionally, TBL allows students to develop their interpersonal communication skills. As part of TBL, students nominate a "team leader" who helps to facilitate the process of scheduling team meetings when students need to further work together outside of

[*] **Joy WAHAWISAN**. PharmD., BCPS. Assistant Professor of Pharmacy Practice. Rangel College of Pharmacy, Texas A&M Health Science Center. Kingsville, TX; & Clinical Pharmacist, Physicians Health Plan, Sparrow Health Systems. Lansing, MI (United States). joy.wahawisan@phpmm.org
Miguel SALAZAR. PhD., PharmD. Associate Professor of Pharmacy Practice. Rangel College of Pharmacy, Texas A&M Health Science Center. Kingsville, TX (United States). msalazar@pharmacy.tamhsc.edu
Robin WALTERS. PharmD. Rangel College of Pharmacy, Texas A&M Health Science Center. Kingsville, TX (United States). rwalters@pharmacy.tamhsc.edu
Fadi M. ALKHATEEB. BSPharm, MBA, PhD. Director of Assessment & Associate Professor. Rangel College of Pharmacy, Texas A&M Health Science Center. Kingsville, TX (United States). alkhateeb@pharmacy.tamhsc.edu
Omar ATTARABEEN. RPh, MS. Department of Pharmaceutical Systems & Policy, West Virginia University. Morgantown, WV (United States). oattarabeen@hsc.wvu.edu

class. Finally, TBL provides students with the opportunity to delegate, manage their time appropriately, and work together in an efficient and quality-based work setting. As pharmacists are required to achieve excellent health outcomes for their patients, student pharmacists need to gain the skills that enable them to guide their patients' treatment plans with evidence-based recommendations in an efficient and timely manner. Therefore, TBL teaches students to improve their work quality, which is key to success in their future careers. For all the above reasons, the Accreditation Council for Pharmacy Education (ACPE) decided to include several elements of TBL instructions starting in the 2007 standards, and continues to do so in the current 2016 standards.[11,12]

Ideally, in a critical thinking discussion course, small groups each with their own clinical facilitator are most conducive to learning.[13]. However, in Colleges of Pharmacy where pharmacy programs are newly established, or in established schools with increased enrollment, the number of faculty and pharmacist preceptors needed to facilitate these courses may not be optimal. With a limited number of clinical faculty available, critical thinking courses in such programs may be led by only one faculty member, coordinating class sizes comprised of 80-90 students. In order to maximize this limited resource in our College, team-based learning was implemented in 2010, and has achieved positive results. To date, students and faculty alike rank the TBL format highly.

Peer evaluation is an integral component of the TBL process. Students are asked to evaluate their team members for three key reasons: (1) to learn how to evaluate other students in a constructive and productive way, (2) to learn how to receive feedback and act upon it, and (3) to incentivize all team members to work together equally. The peer evaluation process we have designed empowers the student to evaluate their peers anonymously and provide comments and constructive feedback to their team members. The peer evaluation is a factor in holding students accountable for their contribution to the team, as a component of their grade is determined by ratings from their team members. Various forms of peer evaluation (though all subjective) have been attempted in previous years with differing degrees of success. Thus, our objective is to test the reliability of our current model of peer evaluation.

The purpose of this study is to evaluate our peer evaluation instrument for reliability. There is limited published literature on the assessment of the reliability of peer evaluation survey tools. This study is designed to assess the usability and reliability of a peer evaluation survey instrument. This may help in better tailoring peer evaluations for future courses. A secondary objective is to assess student perceptions on the change in team dynamic after receiving the results of the first midpoint peer evaluation, as students are asked to assess each other twice every semester. We evaluated students' perceptions of the peer evaluation process with the following criteria: whether the peer evaluation process affects their own behavior, their team members' behavior, or potentially affects the group dynamic after receiving written feedback from the first peer evaluation.

METHODS

In the Rangel College of Pharmacy's (COP) PharmD. Curriculum, "Integrated Pharmacotherapy Sequences (IPT) Rounds and Recitations" is a longitudinal, 4-semester case-based discussion course. Each semester is 1 credit hour, consisting of 45 contact hours. The course is conducted in a team-based learning environment. Students work together in teams of 5 – 7 members to solve problems and discuss patient care. Patient case studies are used as disease management scenarios in order to emphasize critical thinking, problem solving, and decision-making process. Students are responsible for participation in drug use decisions, devising rational pharmacy care plans (therapeutic strategies) and determining optimal drug dosage regimens. Additionally, appropriate parameters for outcome monitoring and assessment techniques are set by students to ensure effective learning about medication safety and efficacy. These workshops are designed to enhance these skills through the application of knowledge learned from medicinal chemistry, pharmacology, therapeutic principles, and pharmacokinetics to the drug management plans of specific diseases and patients. Furthermore, IPT Rounds and Recitations provides an opportunity for students to further develop their clinical skills through specific lab exercises with increasing complexity as the semester progresses. Through these courses, students also hone their skills in discovering and utilizing appropriate literature to help them make appropriate recommendations to various medical professionals.

Teams are randomly assigned through Blackboard (Blackboard Learn™, 2012) at the start of each semester. Participation is mandatory for all students. Student's grades are based upon individual and group work, as well as their peer evaluation. Their final grade is comprised of 45% individual grades, 45% group grades, whereas the remaining 10% comes from the peer evaluation.

Students rate the contributions of each team member twice during the semester through Qualtrics (Qualtrics Survey Software, 2013) and are instructed to assign a score to each team member based on their judgment of his/her quality of work, magnitude of contribution, clinical communication skills, and overall timeliness and responsibility (Online appendix). The first evaluation of the semester is considered "practice". Students receive a completion score on correctly following the instructions, and the results are not factored into their grade. The written feedback is aggregated and distributed back with any identifying information removed, to provide team members an opportunity for improvement. However, assessing individuals' contributions to the team success is difficult, and can become problematic when a significant part of

the final grade is determined by the outcomes of team work. Because peer performance is evaluated twice, the feedback from the initial evaluation gives students a chance to improve their performance.

The results from the second peer evaluation counts as part of the final grade. Points that students receive from the members of their team are averaged to an individual score for every student. Students are instructed not to assign the same grade for all members in their team. As a result, at least one student must receive a score of 9 or lower, and another must receive one score of 11 or higher. Overall, an average score of 10 points is assigned to each team member. For example, in a six-member team, each student has a total of 50 points to distribute among 5 members. Students are instructed to strive to be as fair and precise as possible when assessing the extent of each member's contribution. Additionally, students are requested to provide a rationale for the ratings through comments and constructive feedback. Comments are then compiled along with the average numerical score given by their group members. The subsequent results are reported back to each student anonymously through Blackboard. Earlier versions of the "Rounds and Recitations" course employed different methods of peer evaluation, mainly consisting of a Likert-style scale based evaluation. However, we found that a third of all students gave everyone on their team maximum scores, calling into question the validity of the student assessments.

This study was submitted to the Texas A&M Institutional Review Board (IRB) for approval before project initiation, and the appropriate student consent was obtained. The goal of the study is to evaluate the current peer evaluation model used in the COP "Rounds and Recitation" course series using two separate cohorts. Data was obtained through a retrospective review of the Spring 2013 "PHAR 815 IPT Rounds and Recitation IV" class as well as prospective data from the Spring 2014 "PHAR 815 IPT Rounds and Recitation IV" class. The data was downloaded from Qualtrics by the primary investigator. Qualtrics is a secure, password-protected, web-based survey program available through the COP that enables users to design and run surveys. All identifying data (e.g., student names) were de-identified and re-labeled as "Student 1," "Student 2," for anonymity. Enrollment consisted of all students enrolled in the course at the pre-specified semester (all students in the class of 2014 and 2015). Responses to the peer evaluation survey were analyzed with descriptive statistics to evaluate inter-rater reliability. Assuming a normal distribution, ANOVA was used to test for reliability. Only numerical scores were collected from the peer evaluation; comments that students left for each other were not obtained for research purposes.

For each cohort, students were divided into groups consisting of 7, 6, and 5 students each, such that each student was evaluated by 6, 5, or 4 of his/her peers, respectively. Then, we evaluated the variance of the scores in each group of students by identifying low, medium, and high scores within each group through calculating the mean and standard deviations of the scores. ANOVA was conducted to analyze the differences between score means for every group. Agreement between performance ratings within each group of students was assessed via intra-class correlation coefficient analysis in STATA 12 (College Station, TX) using a one-way random effects model.

For the second objective of this research project, an online survey through Qualtrics was created to assess students' perceptions on team dynamics for the "class of 2015" cohort, conducted after the students had received their results from the first peer evaluation. One of the goals of the peer evaluation process is to positively change or reinforce effective team behavior. A survey was conducted to evaluate the students' perception on the peer evaluation process. Students were asked four questions that were answered in a Likert-style survey response:

1. Whether they believed that the feedback they had received from their peers was accurate.
2. Whether their own behavior changed based on the feedback received.
3. Whether they perceived that their team members' behavior changed in response to the

Table 1. Means and standard deviations of student ratings categorized by low, medium, and high scores in each group

Number of raters per student	Range of sum of scores	Mean of scores	SD	Overall Mean	Overall SD
Class of 2014 Cohort					
6	50-56 57-62 63-70	8.93 9.94 11.0	1.05 0.63 1.30	10.0	1.36
5	42-49 49-50 51-56	9.53 9.97 10.6	0.67 0.51 0.75	10.0	0.79
4	36-38 39-41 42-43	9.25 10.1 10.6	0.79 0.65 0.63	9.93	0.86
Class of 2015 Cohort					
6	53-58 59-62 63-65	9.38 10.0 10.7	0.64 0.56 0.60	10.0	0.79
5	40-48 49-51 52-57	9.13 10.0 10.6	0.92 0.82 0.82	10.0	1.05

SD=Standard deviation

Table 2. Intra-class correlation coefficients between performance ratings

Raters per Student	Number of students	ICC	95%CI	p-value
Class of 2014 Cohort				
6	21	0.795	(0.622-0.905)	<0.001
5	50	0.691	(0.533-0.808)	<0.001
4	15	0.565	(0.055-0.837)	0.018
Class of 2015 Cohort				
6	28	0.810	(0.675-0.901)	<0.001
5	60	0.694	(0.552-0.801)	<0.001
ICC= Intra-class correlation coefficient; 95%CI= 95% confidence interval				

feedback received from the first peer evaluation.

4. Whether the peer evaluation process improved the team's ability to function.

RESULTS

For both the 2014 and 2015 student cohorts, the score means and standard deviations for student groups categorized by low, medium, and high scores for each of the three groups of students are shown below (Table 1). Similar trends were observed in both cohorts. Overall, given the range of scores students could give each other, we found little variation in the standard deviation from the score means in the high, medium, and low scores within each group. Unsurprisingly, the overall means for all three groups were very similar with correspondingly narrow standard deviation widths. In both the 2014 and 2015 cohorts, the intra-class correlation coefficients showed a strong concordance in the ratings that students gave each other, particularly in the larger-sized groups when students were evaluated by six of their peers (2014 cohort ICC=0.795 (95%CI 0.622-0.905; p<0.001); 2015 cohort ICC=0.810 (95%CI 0.675-0.901; p<0.001) (Table 2).

Students' feedback about this evaluation process is shown below (Table 3). The majority of students agreed or somewhat agreed that 1) the feedback they received from their peers was accurate, 2) their own behavior changed based on the feedback received, 3) their team members' behavior changed in response to the feedback received from the first peer evaluation, and 4) the peer evaluation process improved the team's ability to function.

DISCUSSION

There was high concordance among raters using our evaluation instrument methodology as indicated by the both the narrow standard deviations in the scores and the intra-class correlation coefficients. Using a commonly used ICC cutoff value of 0.75 to indicate high agreement between raters, both the large groups in each cohort demonstrate survey instrument reliability.[14] As the intra-class correlation coefficients indicate, using the present methodology, the reliability of scores can be improved by increasing the number of members in each team rating each individual student. Students who are generally stronger performers or contributors to the team are rated as such (i.e. consistently given a higher score by their peers) and students who are not as invested or contribute less to the team have ratings that appropriately reflect their contributions. In other words, through using this survey instrument, students would consistently receive similar scores from each of their team members, whether it be high, low, or moderate, and not widely variable scores as demonstrated by the narrow standard deviation. This finding follows the same trend we found below (Table 3) as the survey results indicated that 72% believe that their peers rated them accurately, and 97% agreed to an extent with their ratings.

Students' attitudes towards the process were assessed through the four-question survey results of the evaluation process. In general, students had a favorable view of the peer evaluation process, and also agreed that their peers rated them accurately. Additionally, more than half of the students believed that they had to modify their own behavior for the better, as well as the behavior of others in their team.

At the end of each semester, students were given the opportunity to provide written feedback regarding the evaluation instrument and its effects in reinforcing or modifying behavior. Comments were largely positive regarding the evaluation process. However, students did address two components that our instrument lacks: a face-to-face discussion, and self-reflection. The lack of face-to-face discussion in the Peer Evaluation process is somewhat inevitable. Many pharmacy students are not comfortable with interpersonal evaluation of their peers.[15] On the other hand,

Table 3. Class of 2015 survey results

Question	Students who responded:			
	Yes N (%)	Somewhat N (%)	Not really N (%)	No N (%)
Did they believe that the feedback they received from their peers to be accurate?	62 (72)	21 (24)	1 (1)	2 (2)
Did their own behavior change based on the feedback received?	33 (38)	33 (38)	12 (14)	8 (9)
Did they perceive their team members' behavior changed in response to the feedback received from the first peer evaluation?	25 (29)	36 (42)	17 (20)	8 (9)
Did the peer evaluation process improve the team's ability to function?	32 (37)	27 (31)	17 (20)	10 (12)

considering the Rounds and Recitation atmosphere through which students interact, a few separate sessions may be conducive to face-to-face evaluation between students. However, an anonymous survey instrument is expected to yield more accurate results, as it is free of the anxiety related with interpersonal interaction. Finally, as pharmacists are required to be good communicators in their profession, student pharmacists may be expected to evaluate their colleagues in a professional sense. But how does one transition the peer evaluation process to an interpersonal, face-to-face discussion? This may be an opportunity for future research.

Students also commented that within their cohort, there were three types of team members: (1) students who contribute, with behavior modification after receiving constructive feedback, (2) students who contribute, with no behavior modification after feedback, and (3) students who minimally contribute, and are not affected by peer feedback into changing behavior. Of these three types of students, this peer evaluation process largely appears to affect the first two types. Therefore, students who have the motivation to improve interpersonal skills, and students who are open to discussion in a TBL setting will benefit the most from a Peer Evaluation process.[16] These claims are subject to interpretation, but the evidence from our data suggests that student-driven motivation strongly affects the results of the survey.

Some students were concerned that they could not give everyone an identical score as they believed that each member was an equal contributor. This sentiment seemed to mostly come up in strong-performing groups with good team dynamics as observed by the facilitators.

A limitation to this study was an inability to formally assess survey validity. Even though we made efforts to evaluate students' responses as critically as possible, truly knowing each individual student's actual contribution to the team discussions and group projects is not easy to be measured. Using an example of a group of 5 students (students A, B, C, D, and E), the survey instrument is designed to measure reliability – do students B, C, D, and E all similarly perceive the amount of work that student A is doing? However, based on our post-assessment survey of the 2015 cohort, the instrument appears to be valid as students generally agreed with their peers' assessment.

Additionally, this study was conducted in one college of pharmacy. Even though we expect the results to apply to other student pharmacists, generalizations may not be made due to variances in geographical area, and may have a different educational background from students of other universities with regard to teaching modalities and peer evaluation. Finally, student demographics were not assessed. It is not known if demographic information (e.g., age, ethnicity, gender) affected the way students evaluated their peers or responded to the survey items.

Overall, this study provides a critical assessment of peer evaluation among student pharmacists. We believe that the results we obtained from our two cohorts validated the reliability of the peer evaluation instrument. Based on our findings, other courses and colleges may be more likely to use this resource.

CONCLUSIONS

An essential part of teamwork accountability is measured through peer evaluation. The peer evaluation instrument, which was addressed in this study, provided a reliable method for peer assessment in a team-based learning setting. Therefore, the results of this study confirmed that students are reliable assessors of their peers. As a result, applying a TBL process, which includes peer evaluation, not only is associated with many benefits to students and facilitators[3,10], but also does not threaten the reliability of individual students' scores that are based on peer evaluations. The findings of this study may encourage academic institutions, which do not have a TBL component in their curricula, to establish a TBL environment and apply the principles of student peer evaluation. Future studies may address how validity can be further assessed and improved for peer evaluation instruments.

Based on students' responses to the survey, the peer evaluation instrument was found to be accurate with regard to students' contributions to the team, and effective in modifying and reinforcing positive teamwork. Therefore, students perceived that their peer evaluation was a valid assessment of their actual contribution to the team effort. This may indicate that pharmacy students are willing to receive feedback from their peers and act upon it, especially when they believe that their peer evaluation can contribute to their final grades. Finally, peer evaluation instrument seem to be constructive in shaping students' academic experience and in helping them advance and excel in a TBL environment with minimal supervision by class facilitators. Future research may focus on how to improve feedback communication channels among students in order to help them correct and improve each other more promptly and efficiently.

CONFLICT OF INTEREST

All authors declare that they have no conflict of interest.

References

1. Garland D. Assessment issues in group work. In Foot HC, Howe CJ, Anderson A, Tolmie A, Warden D, Eds. Group and Interactive Learning. Southampton: Computational Mechanics Publications;1994. p.417-422.

2. Allen RE, Copeland J, Franks AS, Karimi R, McCollum M, Riese DJ 2nd, Lin AY. Team-based learning in US colleges and schools of pharmacy. Am J Pharm Educ. 2013;77(6):115. doi: 10.5688/ajpe776115
3. Ofstad W, Brunner LJ. Team-based learning in pharmacy education. Am J Pharm Educ. 2013;77(4):70. doi: 10.5688/ajpe77470
4. Chung EK, Rhee JA, Baik YH, A OS. The effect of team-based learning in medical ethics education. Med Teach. 2009;31(11):1013-1017. doi: 10.3109/01421590802590553
5. Wiener H, Plass H, Marz R. Team-based learning intensive course format for first-year medical students. Croat Med J. 2009;50(1):69-76.
6. Vasan NS, DeFouw DO, Compton S. A survey of student perceptions of team-based learning in anatomy curriculum: Favorable views unrelated to grades. Anat Sci Educ. 2009;2(4):150-155. doi: 10.1002/ase.91
7. Koles PG, Stolfi A, Borges NJ, Nelson S, Parmelee DX. The impact of team-based learning on medical students' academic performance. Acad Med. 2010;85(11):1739-1745. doi: 10.1097/ACM.0b013e3181f52bed
8. Redwanski J. Incorporating team-based learning in a drug information course covering tertiary literary. Curr Pharm Teach Learn. 2012;4(3):202-206. doi: 10.1016/j.cptl.2012.04.004
9. Carmichael J. Team-based learning enhances performance in introductory biology. J Coll Sci Teach. 2009;38(4):54-61.
10. Fatmi M, Hartling L, Hillier T, Campbell S, Oswald AE. The effectiveness of team-based learning on learning outcomes in health professions education: BEME Guide No. 30. Med Teach. 2013;35(12):e1608-e1624. doi: 10.3109/0142159X.2013.849802
11. Accreditation Council for Pharmacy Education. Accreditation standards and guidelines for the professional program in pharmacy leading to the Doctor of Pharmacy Degree. Guidelines Version 2.0. 2011. Available at: https://www.acpe-accredit.org/pdf/CPE_Standards_Final.pdf (accessed April 1, 2015).
12. Accreditation Council for Pharmacy Education. Accreditation standards and key elements for the professional program in pharmacy leading to the doctor of pharmacy degree, standards 2016. Available at: https://www.acpe-accredit.org/pdf/Standards2016FINAL.pdf (accessed Jan 14, 2016).
13. Nelson M, Allison SD, McCollum M. Luckey SW, Clark DR, Paulsen SM, Malhotra J, Brunner LJ. The Regis Model for pharmacy education: A highly integrated curriculum delivered by Team-Based Learning (TBL). Curr Pharm Teach Learn. 2013;5(6):555-563. doi:10.1016/j.cptl.2013.07.002
14. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standard assessment instruments in psychology. Psychol Assess. 1994;6(4):284-290.
15. 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(8):177. doi: 10.5688/ajpe778177
16. Wahawisan J, Salazar M, Bremick MA. Assessment of a Peer Evaluation Instrument in a Team-Based Learning Course. 115th Annual Meeting of the American Associaton of Colleges of Pharmacy, Grapevine, TX, July 26-30, 2014. Available at: http://www.ajpe.org/doi/pdf/10.5688/ajpe785111 (accessed Jan 14, 2016).