

REVIEW

A Practical Review for Implementing Peer Assessments Within Teams

Anita Yang, PharmD,^a Anna Brown, PharmD,^a Rachel Gilmore, PharmD,^a Adam M. Persky, PhD^{a,b}

^a University of North Carolina at Chapel Hill, Eshelman School of Pharmacy, Chapel Hill, North Carolina

^b Associate Editor, *American Journal of Pharmaceutical Education*, Arlington, Virginia

Corresponding Author: Adam M. Persky, University of North Carolina at Chapel Hill, Eshelman School of Pharmacy, 325 Beard Hall, CB#7569, Chapel Hill, NC 27599. Tel: 919-966-9104. Email: apersky@unc.edu

Submitted June 30, 2021; accepted October 1, 2021; ePublished October 2021

Objective. The objective of this review is to maximize the utility of in administering peer assessments in teamwork settings in professional pharmacy curricula.

Findings. There is a lack of best practices for using peer assessments in the academic setting. The studies reviewed used peer assessments of team for formative and summative assessment, assessing teamwork at limited number of time points to multiple time points; attaching student names to the assessment or being anonymous, and with explanation of why the tool was being used to no explanation.

Conclusion. To get the best use of peer assessments, instructors must define the purpose for their use, explain the purpose of teamwork, orient students to the tool being used, assess teamwork over time and provide feedback, minimize grades associated with the assessment, and use partial anonymity when collecting feedback.

Keywords: peer assessment, collaboration, teamwork, CATME, group work

INTRODUCTION

The use of peer assessments of teamwork (PET) as a measure of performance, contribution, and feedback has been a long-debated topic across many different disciplines. Within education alone, many literature reviews of peer assessments of team and group work exist.¹⁻⁶ Proponents of PETs cite that they help hold individuals accountable for their own contributions to the team, increase engagement among team members, and ultimately contribute to better learning outcomes.⁷ A recent meta-analysis of 55 articles related to PETs found significant evidence between the use of PETs within courses and improvement in academic performance.⁵ Despite these findings and the many perceived benefits of PETs, numerous challenges remain in administering peer assessments, including reliability and fairness.⁸ In addition, there are logistical concerns including the time burden to faculty members, student response and frequency, and grade allotment. Given the increased emphasis on providing team-based health care along with supporting literature demonstrating improved learning outcomes, there exists a need for health care educators to focus on building “team players”.^{9,10} Within pharmacy specifically, the appropriate administration of PETs could serve as a tool to provide feedback to individual students, help them learn and practice team-based principles in a low stake setting, and improve the “soft skills” required to lead a successful career in pharmacy. As such, having consistent guidance regarding PETs and their use across the curriculum would be beneficial. Widespread distribution of techniques and best practices regarding PETs within teams can be used to ensure the purpose of PETs is being achieved.

PETs have been identified as one of the core elements of team-based learning and are already frequently used in pharmacy, health-science education, and many other fields.¹¹ In addition to team-based learning, various forms of learning, such as problem-based learning, collaborative learning, and cooperative learning may warrant the use of PETs. However, despite the ubiquity of these evaluations, there is an unclear consensus among best practices for implementation and assessment.³ There are a wide variety of tools available to implement PETs, yet many institutions continue to develop their own. When working to develop a validated PETs tool, Freeman and colleagues found that four of the main factors that PET schemes differed on were: the use of both self and peer assessment or peer assessment alone; the use of holistic or multiple assessment criteria in assessing teamwork; whether the criteria reflect task or teamwork elements or a combination; and the mathematical formula used to calculate mark weighting factors.¹² Evidence suggests that there are numerous factors that affect the quality of PETs, including the reliability of the instrument, use in different peer to peer interactions, and the stakes of the evaluation.¹³ These factors not only impact the quality of the evaluation, but they can also impact the efficacy of the PET.¹³ Therefore, an inherent limitation of PETs, across all disciplines, is how to create and administer a high quality, reliable PET. The purpose of this review is to suggest best practices for the application of PETs using literature from a variety of discipline.

An Overview of the Benefits and Limitations of PETs

Before we discuss the best practices for the application of PETs, we will briefly review the benefits and limitations of peer-evaluations of teams. There is sufficient evidence suggesting team-based learning, peer-to-peer teaching, and collaborative learning enhances knowledge and critical thinking skills among students.¹⁴⁻¹⁸ For these reasons, many schools of pharmacy have integrated these methods of active learning among their various curricula. An evaluation in 2013 demonstrated that over one-third of schools of pharmacy had implemented components of team-based learning.⁹ It is likely that this number is significantly higher now. In addition, the same faculty that were surveyed indicated that they perceive team-based learning to be more effective than traditional lectures at fostering learning in all 6 domains of Bloom's Taxonomy.⁹ Outside of educational outcomes alone, the push towards value-based health care emphasizes collaboration and the use of interprofessional health care teams, which have been shown to improve a variety of healthcare outcomes and decrease health care costs.¹⁹ The latest standards of the Accreditation Council for Pharmacy Education (ACPE) require that curricula prepare students to be contributing members of health care teams in a variety of settings.²⁰

A review of the literature revealed both positive and negative outcomes of PETs. (Table 1) The first positive outcome is that PETs reduce "social loafing" which is defined as a concept that people are more likely to exert less effort when working collectively as a group compared to performing work individually.^{21, 22} Social loafing can be detrimental as it can reduce productivity, working below individuals' fullest potentials.^{21, 22} Additionally, it can be difficult to identify the contribution of each team member. PETs can reduce social loafing as team members are held accountable for their individual work, which can ultimately increase responsibility and team effectiveness.²³ The more effective the team, the more effective, hopefully, the learning. Second, when students are aware of impending PETs, this can affect their behavior and dialogue when working with their teammates. This could lead to students becoming more cognizant of contributions from their teammates and ensuring they are not underperforming.²⁴ Thirdly, course grades and individual project grades seem to correlate with team functionality, with better performing teams leading to higher team grades and performance.^{7, 25} Finally, the use of PETs also helps students develop skills in providing feedback and potentially conflict resolution.

A considerable factor includes the reliability of the PETs instrument.¹³ Limitations of peer assessments include concerns of quality regarding fairness, reliability, validity, and accuracy.^{1, 26} Because students are more conscious of their upcoming evaluations by their peers, students may spend more time focusing on making a positive impression rather than making meaningful contributions.²⁴ Additionally, bias may present if peer assessments are linked towards students' final grades. Results from Sridarhan and colleagues found that students tended to be overly generous when assessing their peers when their evaluations were incorporated to their final grade, with more pronounced bias in underperforming students.²⁷

Ultimately, despite potential limitations, supporting literature suggests that peer assessments are beneficial and lead to improved outcomes and skills when administered thoughtfully and accurately, which includes implementing safeguards or best practices against the inherent limitations of peer assessments.

Common Tools within the Literature

Due to many researchers and educators customizing their own peer assessment metrics, there is not a consensus on a universal tool to be used in a standardized manner.²⁸ However, a wide variety of published tools exist and we will focus on the most commonly used instruments (Table 2). The Comprehensive Assessment of Team Member Effectiveness, *CATME*, is an instrument that collects behavioral data on teams in five areas research has shown to be important: contributing to the team's work, interacting with teammates, keeping team on track, expecting quality, and having relevant KSAs (knowledge, skills, and abilities).²⁹ *TEAM UP* and *TEAMQ* has been found to be an effective tool that focuses on the 5 domains of teamwork: project planning and management, fostering a team climate, facilitating the contribution of others, managing conflict and contributing to team projects.³⁰ In addition to the overall recommendation and task contribution, *Peer Assess Pro* also highlights the leadership and team processes.

Tools that can be further customized include *SPARKPLUS* and *TEAMMATES*. *SPARKPLUS* is a web-based self and peer assessment tool to encourage students to negotiate the strategy they will work as a team to best achieve their results through equal contributions and improved learning outcomes.^{12, 31} This tool can be customized to select tasks and attributes to be assessed by students' individual work and their peers. *TEAMMATES* provides flexible feedback methods with various visibility control and generates downloadable reports and statistics. Instructors can customize their peer assessment based on multiple choice questions, numerical scale and free-text comments.

The tools provided within this review are by no means an exhaustive list of all available standardized peer assessment instruments. However, these are among some of the most common and validated tools available.

Methods to Decide on Components within Peer-Evaluations of Teams

The most common components of peer assessment tools that can be considered for assessment are listed in Table 3 and fall into 3 categories: Task Contribution and Reliability to the team; Interpersonal Skills; and Leadership. When deciding on what components to include, there are three major ways to determine those components. The first is to use an already developed tool (see Table 2) as these already identify important component pieces. The second is to start with an already available tool (like the ones in Table 2) and customize it based on the needs of the learning environment. Some commercial products do allow customization (eg. SPARKPLUS and TEAMMATES). The third is to establish an evaluation de novo. The advantage of using a previously available instrument is there is minimal development necessary, and the assumption that the tool is valid or reliable, with prior literature to compare performance. The disadvantage is the danger of assuming generalizable validity and reliability to the learners in which the instructor is applying the assessment. For example, if the tool was validated on undergraduate students, it may not necessarily apply to professional students. The issues of validity or reliability may be less important if being used for formative feedback with multiple sample points but become increasingly important for more scholarly pursuits. In contrast, establishing a customized instrument allows students to voice an opinion on their values. This can help with motivation (see self-determination theory) and buy-in as the students help form the rules for classroom management.^{32, 33, 34, 35} Creating an evaluation tool de novo also often minimizes costs, as previously established and validated tools may require subscription or purchase for use.

Administration of Peer assessments

In addition to determining what to assess in a peer assessment, consideration must be given to the timing and frequency of administering peer assessments. Table 4 provides a summary of published literature that provided detail regarding peer assessment administration considerations. Peer assessments with early implementation and multiple evaluation points could assist in alleviating potential barriers associated with peer assessments.²¹ Students who are exposed to peer assessments at the beginning of their teamwork provide the opportunity for teammates to determine their individual role and future contributions for the group.³⁶⁻³⁸ However, peer assessments should not be given solely at the beginning of teamwork, as this could lead to lack of motivation and decrease opportunities for personal development. Therefore, providing multiple evaluation points allows students to recognize areas of improvement necessary to achieve success. Feedback throughout group work fosters open communication reflecting on an individual's role and contributions, along with assessing their peers and group's success overall.³⁶ Ultimately, early introduction to peer assessment along with multiple evaluation checkpoints lead to increased reflection and communication, task motivation, cohesion and can decrease social loafing.

The next issue in implementation is the integrity of peer assessments through anonymity. Studies have found that peers may be hesitant to evaluate each other, and would prefer to remain anonymous.²⁸ Various factors can bias peer assessments such as friendship, peer pressure, ego, age, and self-esteem.^{2, 39} Students should feel reassured that their evaluations are confidential and secure to promote accurate and constructive feedback for themselves and their peers. Advancement in modern technology has simplified the anonymity of administering peer assessment, protecting both identities of assessors and assessees.⁴⁰ However limitations exist within anonymous peer assessments, including encouraging unfavorable behaviors by group members. The lack of necessary confrontation allows students to tolerate their teammates' behaviors with the intention they can penalize underperformers at the final peer assessment.⁴¹

A grading method for peer assessment of group work may be necessary to combat inequitable peer contributions.^{42, 43} Peer assessments can measure group members in a formative and summative manner. In a formative assessment, peer assessments allow individuals to self-reflect on their individual and teammate's work and behaviors along with identifying areas of improvement to ensure their grade accurately reflects their efforts. Formative function can facilitate conflict resolution, team dynamics and overall productivity within a group. Summative peer assessment is performance based as a strategy for instructors to utilize peer assessment information in assigning individual course grades.

A study conducted by Sridharan and colleagues showed that overall, students had the ability to evaluate their peers with accuracy and consistency in an unbiased manner when their feedback does not count towards the final grade.²⁷ For formative assessment, students judged their peers more honestly, but are overly generous when evaluations were attached to a grade. In summative assessment, student evaluations of their peers showed a dramatic inflation and inability to differentiate high contributing students from their counterparts when grades were associated.²⁷ Therefore, it is crucial to outline strategies to mitigate bias such as appropriate policy measures and peer assessment training, incentives and penalties. It is essential to clearly indicate in the syllabus and in lecture how peer assessments will be utilized for individual grades, as students' active participation and meaningful input are pivotal in ensuring validity and reliability.^{42, 44} However, it should be noted that a single numeric value cannot accurately capture the degree of competency for complex skills like teamwork.⁴⁵

Recommendations

Based on the review of literature, we can make several recommendations in administering peer review of teams (Table 5). Peer assessments can be an informative tool that encourages students to be held accountable for their individual and collective contributions to a team. While there is not a universal tool to execute peer assessments, online systems may be used for more accurate peer assessment.⁴⁶ Evaluation tools should be catered to specific goals the course director hopes to achieve through peer assessment. Feedback may be customized to provide timely and insightful evaluation regarding team members' interpersonal and team skills and individual accountability, which can ultimately prevent "social loafing" and promote effectiveness as teams progress through the stages of team formation.¹¹ Michalsen and Fink emphasize that a peer assessment system must be capable of accommodating to different team sizes; accurately reflecting work of team members; and making a significant impact on the course grade.¹¹ Peer assessments can be structured to not only receive constructive feedback, but to also develop skills to provide instructive feedback, both of which will be influential to students' continual progress.¹¹

When previously established peer assessment tools are not used, instructors may consider utilizing a formative structure for evaluations, such as the "keep, start, stop" method or narrative feedback.^{4, 47-50} Crucial to the value of peer assessment activities is adequate orientation and training of students to the tool being utilized. Without thoughtful training in how and why peer assessment is being implemented, the quality of feedback given, and thereafter the impact of the exercise overall, is compromised. Student pharmacists should be trained to be selective with feedback, focusing on one, two, or three most important areas for improvement, rather than diluting the impact of their feedback by listing every detail that can be improved upon.

Whenever peer assessment is utilized, it is important to orient the students to the purpose of the assessments and to the tools they will be using. Throughout courses, peer assessment frequency should be considered, as students use peer assessment tools more often, they become more familiar with the tool to provide feedback more efficiently. Consider the advantages of using peer assessments as formative assessments during important milestones within a given course. Students may gain the advantages of improving their comfort level and skill set with providing feedback throughout the course, which is especially useful if peer assessment is being utilized as a summative assessment at the end of a course. Additionally, students can adjust their performance throughout the course based on peer feedback to maximize their potential.⁷ For these reasons, it is advisable to use the same tool throughout the curriculum. The potential benefits of more frequent peer assessment, however, must be weighed against the time and energy required of faculty and students to coordinate and complete the evaluations. Online tools and software are useful for minimizing time burden as much as possible.

Regarding timing, there is a lack of strong evidence indicating a specific time to conduct peer assessments that is better than any other time. Some studies suggest earlier exposure to peer assessment allows students to acclimate better to the process of both evaluation and overall teamwork.³⁶⁻³⁸ There is, however, no clear indication of when the best time to conduct an initial peer assessment might be, and when it might be "too early" for students to provide their peers with fair and accurate evaluations, versus shallow first impressions. When considering timing, one must decide whether to allot class time to peer assessments or administer the evaluations outside of the classroom time. If the latter route is taken, then an appropriate amount of time between release and due date must be considered, including perhaps the overall workload students may be experiencing at a given time in a semester and exhaustion following submission of a major deliverable.

Anonymity has several advantages in the setting of peer assessments but eliminating anonymity or opting for only partial anonymity may promote responsibility and professionalism. Rather than guarantee anonymity, instructors may consider requiring students to turn evaluations in with their names attached, with the promise of blinding evaluations prior to distributing them to the person being evaluated. This strategy allows for the accuracy and constructivism of anonymous feedback, while promoting professionalism and integrity, as students know the instructors can read what they have written. Other strategies include requiring students to discuss their feedback with their group members, whether the written evaluation was anonymous or not. As health care professionals in training, student pharmacists must learn how to handle difficult conversations. A face-to-face discussion addressing feedback allows students to own their comments but engage in a productive conversation to promote progress and growth individually and as a unit.

Finally, instructors must decide whether incorporating peer assessment scores into students' overall course grades is a worthwhile strategy to reduce social loafing, or a potential risk of compromising the integrity of the evaluations themselves. Potential strategies to address this dilemma include combining formative and summative peer assessments, requiring written justification of scores given to peers, and placing a maximum value on the grade weight of the peer assessment.^{7, 11} By practicing providing feedback in a formative way throughout a semester or project, students can improve their feedback skills and comfortability, as well as adjust their own practices, prior to completing evaluations that count for a grade. Written justifications with concrete examples, though they may introduce more work for both students

and course instructors, minimize risk of falsely inflated peer assessments due to worry over costing a classmate a grade. By capping the potential impact of peer assessment on a student's final grade, students may feel less burdened by that same worry.

CONCLUSION

Overall, the literature is lacking in determination of optimal composition, length, frequency, timing, anonymity, and grading of peer assessments. The overall recommendations summarized in Table 5 aim to initiate the discussion of best practices for peer assessments in professional pharmacy curricula, as the lack of guidance in this space creates an exciting area for further research and experimentation. Further study into each of these areas is needed to develop more concrete recommendations for universal implementation.

With the clear trend in pharmaceutical education moving towards the use of teams in the educational setting, practices must be implemented that promote equal contribution of individuals within teams, and the development of strong teamwork skills. Similarly, as provision of health care continues to become more interdisciplinary and collaborative, pharmacy graduates need to be effective team members upon entry to the workforce. Peer assessments are an important practice to promote accountability, teamwork, and the ever-challenging skills of both giving and receiving constructive feedback.

While a lack of universal guidance regarding peer assessments, particularly in the pharmaceutical education setting, can lead to confusion, it also translates to a broad range of customizability that can allow for maximal utility of peer assessments, when planned thoughtfully. The implementation of goal-centered peer assessments with proper orientation to the tool, strategic timing and frequency, and thoughtful use as a grading mechanism can be impactful in facilitating the development of crucial teamwork skills in student pharmacists.

REFERENCES

1. Dochy F, Segers M, Sluijsmans D. The use of self-, peer and co-assessment in higher education: A review. *Stud High Educ.* 1999;24(3):331-50.
2. Robinson J. Computer assisted peer review. In: Brown S, Bull J, Race P, eds. *Computer-assisted assessment in higher education.* London: Kogan Page; 1999:xi, 205 p.
3. Speyer R, Pilz W, Van Der Kruis J, Brunings JW. Reliability and validity of student peer assessment in medical education: A systematic review. *Med Teach.* 2011;33(11):e572-e85.
4. van de Ridder JMM, McGaghie WC, Stokking KM, ten Cate OTJ. Variables that affect the process and outcome of feedback, relevant for medical training: a meta-review. *Med Educ.* 2015;49(7):658-73.
5. Double KS, McGrane JA, Hopfenbeck TN. The Impact of Peer Assessment on Academic Performance: A Meta-analysis of Control Group Studies. *Educ Psych Rev.* 2020;32(2):481-509.
6. Morbitzer KA, Olsen AA, McLaughlin JE. A mapping review of teamwork training and assessment in pharmacy education. *Am J Pharm Educ.* 2021;85(3):185-95.
7. Mentzer N, Laux D, Zissimopoulos A, Richards KAR. Peer evaluation of team member effectiveness as a formative educational intervention. *J Tech Educ.* 2017;28(2):53-82.
8. Fellenz MR. Toward Fairness in Assessing Student Groupwork: A Protocol for Peer Evaluation of Individual Contributions. *J Manage Educ.* 2006;30(4):570-91.
9. Allen RE, Copeland J, Franks AS, et al. Team-based learning in US colleges and schools of pharmacy. *Am J Pharm Educ.* 2013;77(6):115-.
10. Ofstad W, Brunner LJ. Team-based learning in pharmacy education. *Am J Pharm Educ.* 2013;77(4):70-.
11. 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-.
12. Freeman M, McKenzie J. SPARK, a confidential web-based template for self and peer assessment of student teamwork: benefits of evaluating across different subjects. *Br J Educ Tech.* 2002;33(5):551-69.
13. Norcini JJ. Peer assessment of competence. *Med Educ.* 2003;37(6):539-43.
14. Cortright RN, Collins HL, DiCarlo SE. Peer instruction enhanced meaningful learning: ability to solve novel problems. *Adv Physiol Educ.* 2005;29(2):107-11.
15. Loes CN, Pascarella ET. Collaborative Learning and Critical Thinking: Testing the Link. *J High Educ.* 2017;88(5):726-53.
16. McInerney MJ, Fink LD. Team-Based Learning Enhances Long-Term Retention and Critical Thinking in an Undergraduate Microbial Physiology Course. *J Micro Biol Educ.* 2009;4(1).
17. Silberman D, Carpenter R, Takemoto JK, Coyne L. The impact of team-based learning on the critical thinking skills of pharmacy students. *Curr Pharm Teach Learn.* 2021;13(2):116-21.

18. Zhang J, Chen B. The effect of cooperative learning on critical thinking of nursing students in clinical practicum: A quasi-experimental study. *J Prof Nurse*. 2021;37(1):177-83.
19. Reiss-Brennan B, Brunisholz KD, Dredge C, et al. Association of Integrated Team-Based Care With Health Care Quality, Utilization, and Cost. *JAMA*. 2016;316(8):826-34.
20. Education ACfP. Accreditation standards and key elements for the professional program in pharmacy leading to the doctor of pharmacy degree. Chicago: Accreditation Council for Pharmacy Education; 2016.
21. Brooks CM, Ammons JL. Free Riding in Group Projects and the Effects of Timing, Frequency, and Specificity of Criteria in Peer Assessments. *J Educ Busin*. 2003;78(5):268-72.
22. van Dick R, Tissington PA, Hertel G. Do many hands make light work?: How to overcome social loafing and gain motivation in work teams. *Eur Busin Rev*. 2009;21(3):233-45.
23. Kelley D. Peer evaluation within a team design project. *J Engineer Tech*. 2015;32(1):44-50.
24. Jassawalla AR, Sashittal HC. What Students Think and Do in Classroom Teams When Peer Evaluations Are Highly Consequential: A Two-Stage Study. *Dec Sci J Innov Educ*. 2017;15(2):219-47.
25. Topping K. Peer Assessment between Students in Colleges and Universities. *Rev Educ Res*. 1998;68(3):249-76.
26. Wen ML, Tsai C-C. University students' perceptions of and attitudes toward (online) peer assessment. *High Educ*. 2006;51(1):27-44.
27. Sridharan B, Tai J, Boud D. Does the use of summative peer assessment in collaborative group work inhibit good judgement? *High Educ*. 2019;77(5):853-70.
28. Loughry ML, Ohland MW, DeWayne Moore D. Development of a Theory-Based Assessment of Team Member Effectiveness. *Educ Psychol Measure*. 2007;67(3):505-24.
29. Ohland MW, Loughry ML, Woehr DJ, and colleagues. The Comprehensive Assessment of Team Member Effectiveness: Development of a Behaviorally Anchored Rating Scale for Self- and Peer Evaluation. *Acad Manage Learn Educ*. 2012;11(4):609-30.
30. Slootweg IA, Lombarts KMJM, Boerebach BCM, Heineman MJ, Scherpbier AJJA, van der Vleuten CPM. Development and validation of an instrument for measuring the quality of teamwork in teaching teams in postgraduate medical training (TeamQ). *PloS one*. 2014;9(11):e112805-e.
31. Willey K, Gardner A. Improving self- and peer assessment processes with technology. *Camp Wide Info Sys*. 2009;26(5):379-99.
32. Deci EL, Ryan RM. The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychol Inq*. 2000;11(4):227-68.
33. Gagné M, Deci EL. Self-determination theory and work motivation. *J Org Behav*. 2005;26(4):331-62.
34. Rivers SE, Brackett MA, Reyes MR, Elbertson NA, Salovey P. Improving the Social and Emotional Climate of Classrooms: A Clustered Randomized Controlled Trial Testing the RULER Approach. *Prevent Sci*. 2013;14(1):77-87.
35. Schiefele U. Classroom management and mastery-oriented instruction as mediators of the effects of teacher motivation on student motivation. *Teach Teacher Educ*. 2017;64:115-26.
36. Druskat VU, Wolff SB. Effects and Timing of Developmental Peer Appraisals in Self-Managing Work Groups. *J Appl Psychol*. 1999;84(1):58-74.
37. Reilly RR, Smither JW, Vasilopoulos NL. A longitudinal study of upward feedback. *Person Psychol*. 1996;49(3):599-612.
38. Smither JW, London M, Vasilopoulos NL, Reilly RR, Millsap RE, Salvemini NAT. An examination of the effects of an upward feedback program over time. *Person Psychol*. 1995;48(1):1-34.
39. Raes A, Vanderhoven E, Schellens T. Increasing anonymity in peer assessment by using classroom response technology within face-to-face higher education. *Stud High Educ*. 2015;40(1):178-93.
40. Li L. The role of anonymity in peer assessment. *Assess Eval Higher Educ*. 2017;42(4):645-56.
41. Bacon DR, Stewart KA, Silver WS. Lessons from the Best and Worst Student Team Experiences: How a Teacher can make the Difference. *J Manage Educ*. 1999;23(5):467-88.
42. Chen Y, Lou H. Students' Perceptions of Peer Evaluation: An Expectancy Perspective. *J Educ Busin*. 2004;79(5):275-82.
43. Williams DL, Beard JD, Rymer J. Team Projects: Achieving their Full Potential. *J Market Educ*. 1991;13(2):45-53.
44. Cederblom D, Lounsbury JW. An Investigation of User Acceptance of Peer Evaluations. *Person Psychol*. 1980;33(3):567.

45. Van Der Vleuten CPM, Schuwirth LWT, Driessen EW, Govaerts MJB, Heeneman S. Twelve Tips for programmatic assessment. *Med Teach*. 2015;37(7):641-6.
46. Smith AR. Managing the Team Project Process: Helpful Hints and Tools to Ease the Workload without Sacrificing Learning Objectives. *E-J Bus Educ Scholar Teach*. 2018;12(1):73-87.
47. Pelgrim EAM, Kramer AWM, Mokkink HGA, van der Vleuten CPM. The process of feedback in workplace-based assessment: organisation, delivery, continuity. *Med Educ*. 2012;46(6):604-12.
48. Pelgrim EAM, Kramer AWM, Mokkink HGA, Vleuten CPMvd. Quality of written narrative feedback and reflection in a modified mini-clinical evaluation exercise: an observational study. *BMC Med Educ*. 2012;12(1):97-100.
49. Ramani S, Könings KD, Ginsburg S, van der Vleuten CPM. Twelve tips to promote a feedback culture with a growth mind-set: Swinging the feedback pendulum from recipes to relationships. *Med Teach*. 2019;41(6):625-31.
50. Delong T. Three Questions for Effective Feedback *Harv Busin Rev*. Aug 4 2011.
51. Levine RE, Kelly PA, Karakoc T, Haidet P. Peer Evaluation in a Clinical Clerkship: Students' Attitudes, Experiences, and Correlations With Traditional Assessments. *Acad Psychiat*. 2007;31(1):19-24.

AJPE
Accepted Draft

Table 1. Evaluation as Reported in the Literature

Outcome	Literature Summary
Positive Outcomes	<ul style="list-style-type: none"> ● Reduced “social loafing” - individual responsibility and team efficacy encouraged²³ ● Improved course and individual project grades⁷ ● More “student-centered” learning environment⁷ ● Enhanced level of care, impression management behaviors, perception of contribution from teammates²⁴ ● Positive formative effects on student achievement and attitudes²⁵
Negative Outcomes and Limitations	<ul style="list-style-type: none"> ● Time intensive for both students and faculty¹¹ ● Highly variable quality - depends on instrument reliability, peer to peer interactions, and stakes of evaluations¹³ ● Questionable fairness, reliability, validity, accuracy^{1, 26} ● Priority placement on impression management rather than meaningful contribution²⁴ ● As summative assessment, can inhibit good judgment (students tend to be overly generous)²⁷
Neutral Outcomes	<ul style="list-style-type: none"> ● Evaluation scores modestly predict student performance on other measures (quizzes, standardized tests, etc.)⁵¹

Table 2. Common Available Peer Assessment Tools

Available Tools	Synopsis of Tool
CATME	<p>Web-based instrument that collects data on team-member effectiveness in 5 areas:</p> <ul style="list-style-type: none"> ● Contributing to team’s work ● Interacting with teammates ● Keeping team on track ● Expecting quality ● Having relevant knowledge, skills and abilities
SPARKPLUS	<p>Customizable web-based self and peer assessment kit</p> <ul style="list-style-type: none"> ● Allows academics flexibility to choose/create specific targeted criteria to allow any task/attribute development to be assessed <ul style="list-style-type: none"> ○ Professional skill examples: enthusiasm & participation, team organization, idea contribution, problem-solving, efficiency, conflict management, constructive feedback, reliability ○ Task-specific performance examples: Evaluating performance by breaking down assignment/semester into smaller parts ● Allows students to self and peer assess individual work and improve their judgement through benchmarking exercises ● Automates data collection, collation, calculation, and distribution of feedback of results ● Anonymous written feedback to their peers
TEAM UP & TEAMQ	<p>Assess 5 domains of teamwork skills</p> <ul style="list-style-type: none"> ● Project planning and management

- Fostering a team climate
- Facilitating the contribution of others
- Managing conflict
- Contributing to team project

TEAMMATES

Customizable, automatically-generated components

- Member contribution
- Comments about their own contribution
- Team dynamics
- Feedback to each teammate

Peer Assess Pro

- Overall recommendation
- Task contribution
 - Initiative
 - Attendance
 - Contribution
 - Professionalism
 - Ideas and learning
- Leadership & team processes
 - Focus and task allocation
 - Encourages contribution
 - Listens and welcomes
 - Conflict management and harmony
 - Chairmanship
- Developmental feedback
 - Qualitative, highs/lows
- Teacher advice

Table 3. Common Components of Peer Evaluation Instruments

Component of Peer assessment	Metrics within Component
Task Contribution & Reliability	<ul style="list-style-type: none"> ● Percent of work done ● Relevant knowledge, skills, abilities ● Contribution of ideas, innovation ● Problem solving & constructive feedback ● Attendance & punctuality ● Quality control ● Efficiency ● Focus
Interpersonal Skills	<ul style="list-style-type: none"> ● Teammate interaction ● Enthusiasm ● Participation ● Conflict management and resolution ● Fostering a team climate ● Harmony ● Listening & welcoming
Leadership	<ul style="list-style-type: none"> ● Keeping team on track ● Encouraging excellence ● Facilitation of others' contributions ● Task planning and management

Summative Measures

- Chairmanship
 - Overall score
 - Overall recommendation
 - Qualitative, developmental feedback (strengths, weaknesses)
-

AJPE
Accepted Draft

Table 4. Selected Studies Demonstrating the Variability in Peer-Assessments of Teams

Reference	Population, Frequency and Timing	Tool Utilized	Anonymity
21	Students in a required, introductory, cross- disciplinary business course. Students were evaluated every 4 weeks at the end of modules following team project and presentation, 3 times during the semester	Custom peer assessment tool determined weight of overall group score allotted to each student. Tool consisting of attendance, punctuality, equal workload, enthusiasm, and attitude.	Anonymous when shared with group members
44	Full-time faculty of a small university evaluated annually at an unspecified timing	Custom tool consisting of unweighted performance criteria: professional development, cooperation with colleagues, teaching effectiveness, contribution to organizational objectives considered in annual evaluations.	Unspecified
36	Students in a required introductory organizational behavior course. Students were evaluated on a questionnaire 2 weeks prior to the peer appraisal, 1 week prior to appraisal, immediately after appraisal, and three class periods after appraisal. Peer appraisals were performed 6-7 weeks after group formation, variable timing relative to group presentation.	Ungraded conglomerate of previously validated scales measuring open communication, task motivation, group viability, group cohesion, and satisfaction with group, social loafing.	Anonymous
23	Students in a freshman level, required engineering graphics course evaluated once following completion of group project	Previously developed ungraded instrument designed to assess work-related categories (quality of work, quantity of participation, timeliness, level of work), and measure student contribution to the project	Unspecified
51	Students participating in psychiatry 6-week clinical clerkship and engaging in team-based learning evaluated during each clerkship cycle. Unclear number of iterations.	Each student was given a supply of 10 points per teammate and instructed to assign the points to the other members based on how they felt the members had contributed to their learning and/or group's performance. Six out of eight student groups chose not to have evaluation count toward grade.	Anonymous
40	Pre-service teachers enrolled in a technology application course evaluated	Self and Peer Assessment Tool using assignment rubric	Study assessed both

once following drafting of assignment and attending peer assessment training. After receiving feedback, students revised and submitted the assignment.

to evaluate other students' performances on a specific deliverable. Quality of peer review was considered for class participation grade.

anonymous and named evaluations

7

Students in freshman level design thinking course evaluated once during each of the three major project deliverables and twice during the early stages of the final project. Evaluations were timed such that students had the opportunity to receive the feedback prior to engaging in the next main deliverable.

CATME tool was utilized, and students were trained on how to give and receive feedback and interpret results. Evaluations were not incorporated into student grades.

Unspecified

37

Corporate managers being rated by subordinates evaluated four times of an upward feedback program.

33 behavioral statements reflecting boss/ subordinate relationships across 9 different performance categories. Unspecified if evaluations were considered in performance evaluations.

Table 5. Recommendations for Implementing Peer Assessments of Teams

Recommendation

Orientation

- Define the purpose of peer-evaluation
- Explain purpose of teamwork and set expectation
- Orient students to the tool

Implementation

- Have multiple assessments over time
- Use the same tool throughout the curriculum
- Provide feedback
- Minimize "grading"
- Use partial anonymity