

Approaches to Peer Evaluation: Pro's and Con's of Various Methods

Rick Goedde, St. Olaf College, goedde@stolaf.edu

Jim Sibley, University of British Columbia, jim.sibley@ubc.ca

Peer Evaluation is an essential component of the Team-Based Learning Model. Ensuring that you develop a peer evaluation method that is both fair and “defensible” is one of the keys to a successful TBL course. Some students may express concern when they are to be evaluated by their peers. It is important to stress the value of peer feedback, and how common peer feedback is in the modern workplace.

In the context of TBL, peer evaluation has teammates assessing teammates in order to “measure” the “helping behaviors” and to develop a peer evaluation scores. Peer evaluation does not measure how much a student knows, but how his or her contribution to team cohesion and productivity is perceived by their teammates.

Peer evaluation can have numerous benefits:

1. By motivating students to strive for excellent performance
2. By giving comfort to “stronger” students who fear that they will be doing most of the work
3. By providing the instructor with a more accurate way of assigning individual grades at the end of the course (students know their teammates’ contributions better than the instructor does).
4. By helping students develop their feedback and evaluation skills

Koles Method:

This method includes both a quantitative and qualitative assessment. The strength of this approach is that both the qualitative and quantitative data is generated, and that a student’s final peer evaluation score is both informed by their work across the semester (quantitative component) and the quality of feedback they give their peers (qualitative component). This method helps students develop their feedback skills.

In the quantitative component, each student is rated on 12 competencies using a likert scale (never, sometimes, often, always). There are twelve competencies divided into three areas: cooperative learning skills, self-directed learning, and interpersonal skills. Points are generated from the compiled likert results for each student.

In the qualitative component, each student answers two questions for each teammate:

What is the single most valuable contribution this person makes to your team?

What is the single most important thing this person could do to more effectively help your team?

The quality of the feedback given is evaluated by the instructor using specific criteria, and up to 40 points is given to the student who made the comments. The scores from the quantitative assessment (60%) and the qualitative assessment (40%) are added together to get the final peer feedback score. Students must complete the quantitative and qualitative sections on all of their teammates to receive their own peer evaluation score.

Pros and cons: The benefit of this method is that your peer evaluation score both depends on the quality of your performance as judged by your peers, and the “helpful” quality of the feedback you give your teammates.

Michaelsen Method:

In this an anonymous method where the peer evaluation makes up a fixed portion of the final grade (5 to 10% typically). The peer evaluation scores let the instructor assign a fair portion of this course grade component after review the evaluations. Often there is an enforced score differential (for example in a 10 point per student rating - not everyone can get a 10 - at least one 11 and one 9). Students typically are required to provide reasons for the highest ratings and lowest ratings.

Pros and cons: This method can be beneficial due to the enforced score differential, but some students complain about having to give scores below 10. Without a differential there can be a tendency for students to assign 10’s across the board in the hopes of maximizing evaluation score for everyone. In schools where a lower limit is not used (i.e. no scores less than 7 out of 10 allowed), some students have noticed that giving one student consistently zero can substantially boost other teammates scores. (Ostafichuk, personal communication)

Combination Method:

A number of practitioners stipulate that peer evaluation score can only RAISE a student’s grade. Although this addresses a discomfort that students may have, it can lead to an increase in “social loafing” since there is a reduced penalty for not contributing to team success.

Fink Method:

This is another anonymous method where the evaluation score is used as a multiplier with the team portion of the course grade. This can be very useful in courses that have a larger team component grade (i.e. project courses). In this method it is possible for a team member to get higher or lower mark than the team product grade - depending on the peer evaluation. For example, one hard working student may get 90% on the report (that was originally marked at 80%) and another student that has contributed less may get 70%. This method can use simple (divide the money) evaluations or rubric-based evaluations. Rubric-based evaluations are often perceived to be fairer by students, as they seem to be more comfortable assigning lower scores, when they can justify them (Ostafichuk, personal communication).

A number of TBL practitioners are now recommending that students develop their own performance criteria, by which everyone will be judged. “Student value that which they create” (Lane, 2007). An activity is added to the start of the semester, where student work in their teams to develop the criteria (the whole class must agree on the same criteria). This can increase student buy-in to peer evaluation.

Pros and cons: Unlike the Michaelsen Method, this method allows students to give the same score to each teammate, which may lead to more student satisfaction. A potential problem is that sometimes students underestimate the degree to which they can affect a teammate’s final grade with relatively small changes in the peer evaluation scores.

Texas Tech Method:

This is another likert scale method where students evaluate their teammates on twelve criteria, including Promptness/Reliability, Humility, and Preparation for Learning Activities. Students are ranked on a five-point scale in which 1 is too little, 5 is too much, and 3 is considered an ideal score. Comments are required for scores of 1 or 5 but are otherwise optional. Scores are not counted when given as formative, but are included in the course grade at the end of the semester. All the data are collected electronically and fed back to the students.

Pros and cons: This method has similar pros and cons as described in the Koles Method above, but the qualitative component of this method is not evaluated with the same thoroughness as the Koles Method.

UT Austin Method:

Students are asked to provide one thing they appreciate about a teammate and one thing they request. Students are given an anonymous compilation of their teammates appreciations and requests. This method does not impact grades directly, but can be used when student request an extra credit bump at the end of semester when they are shy of the-next-grade-up. The possibility of the “need” to use evaluations to get the “bump” might keep them engaged all semester long. (Sweet, 2007)

Wisdom of ListServe

Peer Evaluation is one of the most common topics discussed on the listserve. There is no one size fits all solution for peer evaluation, but there are many things you can do to be more successful.

- Many TBL practitioners use “Making Feedback Helpful” by Michaelsen and Schultheiss
- Students should be introduced to the logistics of peer evaluation early in the semester.
- Students should be reminded of purpose and your rationale’s for using peer evaluations
- Students should be made aware of similar peer assessment processes that are being incorporated into many work place environments

- We should be made it clear how the evaluation score could affect a student’s individual final mark.
- Most TBL practitioners do a formative evaluation part way through the semester (typically at end of first month)
- Using student generated performance criteria can increase student buy-in
- Using student Grade Weighting Activities can increase student buy-in
- Many practitioners accept student efforts to collude to eliminate enforced score differentials as reasonable and accept it as “team work”

Essential Reading

- Michaelsen, Parmelee, McMahon & Levine, (eds.), Team-Based Learning for Health Professions Education, Sterling, Va., Stylus 2008
- Michaelsen, Parmelee,, Sweet, (eds.) Team-Based Learning: Small Group Learning’s Next Big Step New Directions for Teaching and Learning, Volume 2008, Issue 116
- Lane, D.R., Peer evaluation in TBL: A student-developed approach to the assessment of peer evaluation in medical education. (2007) PODCast Proceedings TBL 2007 Conference, available at <http://tbl.apsc.ubc.ca/conferences/2007/#a> (viewed on February 22, 2011)
- Michaelsen, L. K., and Schultheiss, E. E. “Making Feedback Helpful.” Journal of Management Education, 1989, 13, 109-113.