5

This chapter discusses the merits of peer assessment and evaluation; the instructional guidelines, issues, and considerations for their use in the TBL classroom, and a brief description of peer assessment methods and information on how to access forms.

Peer Assessment and Evaluation in Team-Based Learning

Christina M. Cestone, Ruth E. Levine, Derek R. Lane

In contrast to traditional courses, in which students are accountable only to the instructor, effective implementation of any group-based instructional format, including team-based learning (TBL), requires that students be accountable to both the instructor and their peers. Unfortunately, some instructors resist using groups because of concerns about using peer evaluations or poorly designed group assignments and grading systems that in effect reward and encourage social loafing. Furthermore, their concerns seem as equally focused on unfairly raising the grades of poor students (who may be carried along by hard-working members) as they are about the potential of penalizing hard-working students (who may receive a lower course grade because they were randomly assigned to a poorly performing group).

With TBL, however, these concerns are largely eliminated by using a comprehensive feedback and grading system that ensures individual student accountability to both the instructor and peers by using a grading system that has an individual performance component, a team performance component, and a peer evaluation component (Michaelsen, 1992; Michaelsen, Knight, and Fink, 2004; Chapter One, this volume). Individual student grade incentives for accountability in TBL courses are derived from four separate but interrelated sources: individual preparation (scores on the individual readiness assurance test); collective preparation (scores on the team readiness assurance test); how well knowledge is applied by individuals

NEW DIRECTIONS FOR TEACHING AND LEARNING, no. 116, Winter 2008 © Wiley Periodicals, Inc.
Published online in Wiley InterScience (www.interscience.wiley.com) • DOI: 10.1002/tl.334

within a team, as a team, to solve progressively difficult problems (application exercise scores); and contributions to interpersonal group dynamics, team maintenance, cohesion, and team productivity (peer evaluation scores).

Peer assessment and evaluation are essential elements of TBL for two reasons. First, peer assessment provides formative (process) information to help individual students improve team performance over time and develop the interpersonal and team skills essential for their future success. Second, peer evaluation scores provide summative (outcome) data to the instructor that can be used to ensure fairness in grading by incorporating an assessment of each member's contributions to the success of their teams and make judgments about it (Falchikov and Goldfinch, 2000; Topping, Smith, Swanson, and Elliot, 2000; Topping, 2005).

This chapter discusses the pedagogical merits of peer assessment and evaluation as mechanisms for enhancing student accountability and presents the instructional guidelines, issues, and considerations for use in the TBL classroom. It concludes with a brief description of peer assessment methods and information about how to access and use the corresponding forms.

Pedagogical Merits of Peer Assessment and Evaluation

TBL is a unique pedagogical strategy that facilitates the development of a variety of skills relative to future professional pursuits, including individual accountability, problem solving, interpersonal communication, teamwork, and organizational skills (Brindley and Scoffield, 1998; Boud, Cohen, and Sampson, 1999; Butcher et al., cited in Cheng and Warren, 2000; Lane, 2007; Levine, 2008). TBL provides students with multiple and varied opportunities to develop these skills as they engage in team activities including team readiness assurance tests (tRATs), group preparation of written appeals, and application-oriented problem-solving activities and exams (Michaelsen, Knight, and Fink, 2004).

Students also receive a number of benefits from engaging in the process of peer assessment and evaluation. For example, judgment by peers seems to provide a more significant motivator to produce high-quality work than does the assessment of a single instructor (Searby and Ewers, 1997). As a result, when peers are accountable to each other, the time spent comparing work and discussing ideas and concepts in teams is more productive. Put simply, students learn from the cognitive processes of their peers (Brindley and Scoffield, 1998). Furthermore, when students become assessors, they are required to show a more thoughtful understanding of the processes involved in the activity (Searby and Ewers, 1997).

In their meta-analysis of peer assessment studies, Dochy, Segers, and Sluijmans (1999) identified several positive effects of using various peer, self, and coassessment processes to improve the quality of learning. Two of the most prominent were increased confidence in one's performance and increased quality in the learning output. Peer assessment and evaluation also

serve to help students take more control over their learning through development of critical analysis of the work of others (Searby and Ewers, 1997). Repeated reading of peer writing, for example, reinforces analysis of one's own work processes and makes the learning more visible to the learner (Topping, Smith, Swanson, and Elliot, 2000). Similarly, music composition students who were responsible for analyzing the compositions of other students were better able to recognize the components of a quality composition against which they could hold their own work (Searby and Ewers, 1997).

Peer assessment is also useful for helping students to scrutinize the purposes and objectives of a course (Smith, Swanson, and Elliot, 2000). When conducted prior to the end of the semester, peer evaluation enabled earlier identification of misunderstandings or gaps in thinking (Searby and Ewers, 1997) that could then be addressed by targeted discussion of specific course content. Finally, peer assessment and evaluation are essential to a comprehensive grading process in TBL because team members are typically the only ones who have enough information to accurately assess one another's contributions (Michaelsen, 2004).

Student Perceptions of Peer Assessment and Evaluation

An examination of the literature across disciplines indicates a range of differences in terms of learner acceptance of peer evaluation. Several studies have demonstrated positive correlations regarding faculty evaluations and written exam performance (Levine, 2008). In some studies, learners expressed satisfaction with and believed they benefited from peer evaluation (Gatfield, 1999; Paswan and Gollakota, 2004), while in others, learners resisted the process. Learners who were accepting of the process believed that the quality of their work improved as a result of the feedback they received (Brindley and Scoffield, 1998; Dochy, Segers, and Sluijmans, 1999; Topping, Smith, Swanson, and Elliot, 2000). Students who disliked peer evaluation believed that the process interfered with the relationships with their fellow learners (Levine, 2008). While some students experienced a sense of socio-emotional discomfort (Topping, Smith, Swanson, and Elliot, 2000), they still perceived an increase in personal motivation as a result of their active participation in peer assessment (Brindley and Scoffield, 1998), including a heightened sense of engagement and concentration. Peer assessment offers students an opportunity to compare their work to that of others. Brindley and Scoffield's study (1998) supports the student view that there is more ownership of the learning experience.

While peer evaluation has the potential to provide valuable feedback to learners, it can foster a classroom environment of distrust and high competitiveness when implemented in a clumsy fashion (Levine, 2008). Students need high-quality and effective peer evaluation to feel comfortable that their teammates are contributing their fair share of the group work

(Levine, 2008). When done well, engaging students in peer assessment and feedback can be a valuable exercise in self-development (Brindley and Scoffield, 1998).

Knowing that they will receive grades or points from peer assessments is a powerful incentive for students to prepare for and participate in the group work of the course (Michaelsen, 1992). However, the peer evaluation process also may reduce student motivation to participate unless its use is clearly communicated and aligned with students' expectations and values for its use (Chen and Lou, 2004). Peer assessment is not a set prescriptive process, but rather one that may take time to develop and may also change over time depending on the course content, class size, the curriculum, the university culture, and the students themselves.

Guidelines for Implementing Peer Assessment and Evaluation

Michaelsen (1992) strongly recommended that instructors set student expectations early and consider collaboratively determining assessment dimensions such as the weight each will contribute to final grades and the frequency with which they are administered. Peer assessment and evaluation should also be aligned with grading policies, classroom procedures, and group activities (Michaelsen, Knight, and Fink, 2004), and instructors should offer resources for students on how to provide constructive evaluation and feedback.

Setting Expectations. The process of involving students early offers the opportunity to determine what dimensions and assessment criteria are relevant to the activities and course objectives of the TBL classroom. This is generally done as soon as students understand how team-based learning works, usually no later than the third class (Michaelsen, Knight, and Fink, 2004). The assessment literature provides a thorough examination of the dimensions and criteria measured by peer assessments in TBL, and forms that can be accessed via the Web site www.teambasedlearning.org/ndtl offer other examples as well. Typically criteria include group process and individual task and individual group behavior contributions such as cooperation, flexibility, dependability, attendance, attitude, respect for team members, preparedness, initiative, leadership, communication, and decision making (Dominick, Reilly, and McGourty, 1997; Schelf-hout, Dochy, and Janssens, 2004; Thackeray and Wheeler, 2006).

Involving Students in Designing Instruments and Procedures. Lane (2007) describes the advantage of allowing peer assessment instruments and procedures to be developed by engineering students as opposed to faculty, which is similar to the early team-building exercise Michaelsen used (1992, 2004) when students engage in setting their own grade weights. Michaelsen identifies three considerations for instructors who are considering how much "grade weight" (p. 113) to give to their peer assessment. First, peer evaluations should be weighted so that students take the process seriously

in terms of impact on their final individual grade. Next, the instructor must be comfortable with administration of the instrument. Third, it should address student concerns for fairness and equity.

When students develop their own assessments, they become invested in the outcomes of their evaluation efforts, resulting in greater ownership for the assessment criteria and associated feedback. Put simply, students support that which they help to create. Falchikov and Goldfinch (2000) suggest that student familiarity with and ownership of assessment criteria also tend to foster peer assessment validity and recommend student involvement in their determination.

Using Periodic Formative Assessments. Using periodic formative assessments has a number of advantages. When used at the end of learning modules, they help assuage student concerns about equity issues (Gueldenzoph and May, 2002; Haberyan, 2007; Levine, 2008). In addition, using periodic formative assessments allows students the opportunity to develop their skills at giving assessments before doing a final summative evaluation (Levine, 2008). Periodic assessments also promote team effectiveness and enable students to improve their own skills. Students who are unaware that their behaviors are unacceptable to team members learn early to change behaviors that may be unproductive or disruptive to the team (Michaelsen, Knight, and Fink, 2004). Some additional suggestions for the process include providing a mechanism for giving anonymous or confidential comments, making sure that members give each other both positive and negative comments, and giving later feedback the most weight (Michaelsen, 2004). Depending on the environment, peer evaluations may be used at the midpoint and at the end of a course. Implementing too many rounds of peer evaluation, however, may have drawbacks. These include disruption to team development and validating the role of a dominant participant early in the semester (Michaelsen, Knight, and Fink, 2004).

Preparing Learners for Feedback. It is important to prepare learners before asking them to participate in peer evaluation. It should not be assumed that students understand how to perform peer evaluation or that they have been exposed to it in other course work. In demonstrating how peer assessment is done in a graduate educational psychology writing course, for example, Topping, Smith, Swanson, and Elliot (2000) reported using a demonstration of an instructor conducting a critical analysis of their own peer-reviewed journal publication to highlight areas for improvement as a means of providing a model for students to follow in conducting their own peer assessments. Students who are intellectually capable but perhaps socially unskilled can learn through exposure to feedback from their peers who have similar outcomes at stake (Michaelsen, 1992).

Giving appropriate and constructive feedback is a skill that takes time and instruction to do well. Early in the semester or on the evaluation form, a verbal or brief written explanation may be provided to students (Levine, 2008). This offers students a guide to follow, which may ease any concerns

on what feedback should be gathered and how it should be structured to be most effective. It may be useful to communicate the hallmarks of effective feedback so students understand what is expected and to ease the social discomfort that may come with the process. Characteristics of effective feedback are addressed at the conclusion of this chapter.

Anonymous Versus "Owned" Assessments and Evaluations. Although there is universal agreement among TBL users that the instructor needs to know who is saying what to whom, there is considerable debate about how much students should know about the source of the scores and comments they receive from their peers. On one hand, students may be more honest when they know their peers will not know which team member offered the feedback. On the other hand, when students do not own the feedback they give (that is, the receiver does not know who gave it), they may provide harsher criticisms and evaluations (Lane, 2007) and as a result have a negative impact on the relationships between team members. If feedback is done correctly, either approach can be successful. Understanding your environment, communicating with students in the process, and being consistent in the administration of peer assessments are key. These suggestions seek to remedy student concerns about fairness, remove the mystery surrounding its contribution to the final course grade, and set expectations about the student and teacher role in the TBL environment.

Customizing the Process. Every educational environment is different, so the use of a peer evaluation instrument for one setting might not be appropriate for another institution or discipline. In many respects, peer review is best received in an environment in which there is a culture of professionalism and a minimum amount of competition and mistrust. The more courses that promote and encourage peer review, the better students will accept it and use it constructively (Levine, 2008).

Part of determining what is right for the particular environment includes determining the frequency with which the peer assessment process will be used. Brooks and Ammons (2003) found that in a cross-disciplinary course with modules on accounting, marketing, and management, implementing peer assessments every four weeks, for a total of three times during the semester, reduced the variation of rating among students. In a human learning undergraduate course, the instructor uses a formative assessment at midsemester and a summative peer evaluation at the end. Levine (2008) also advocates a similar process of two times during the semester for students in a clinical graduate course (Levine and others, 2004).

Assessment Instruments and Approaches

For instructors who lack the time or interest to design a peer assessment in its entirety, excellent instruments are available. We recommend considering five approaches. Two of the approaches were developed by Michaelsen and Fink and are described in detail in Michaelsen, Knight, and Fink (2004);

the related forms are contained there as well. Briefly, in the Michaelsen method (2002), students are expected to assign teammates a score based on the extent to which they believe their teammates contributed to the overall team performance. For example, in a six-person team, fifty points are given to each student to divide among five team members (self-excluded), with a minimum possible score of seven, average of ten, and maximum of thirteen (Levine, 2008). The overall score for an individual is then calculated by summing the scores received from each teammate. Students also have an opportunity to include qualitative comments. This method requires that students make distinctions among peer performances; not everyone can receive a ten.

In the Fink method, students are given one hundred points and prompted to divide them among team members based on their degree of contributions. All members then get a peer score that is the sum of the points they are awarded by each team members and then this total is multiplied by the their mean readiness assurance test score (or another group score) to come up with an adjusted group score. Students are also prompted to provide qualitative feedback with justification for the number of points that were assigned. This method differs from the Michaelsen method in that students may assign all one hundred points to each peer; there is no required differentiation of points.

A third approach, developed by Paul Koles for use with year-long medical student teams, is outlined in more detail in Levine (2008). It includes both a comprehensive quantitative feedback section capturing ratings on cooperative learning skills, self-directed learning, and interpersonal skills and qualitative questions. Qualitative questions probe the most valuable contributions a person makes to the team and the most important thing a person could do to more effectively help the team (Levine, 2008). Feedback is anonymous.

A fourth approach, outlined in Chapter Six in this volume, involves using qualitative data (peer comments given and received) as a "difference maker" for students whose grades fall on a borderline. Finally, Lane (2007) describes an approach that involves students in creating the instruments and procedures that they will use for collecting quantitative and qualitative peer evaluation data that is then used to provide feedback and grading input for the members of their team. (All five of these approaches and the forms that support their implementation are available online by clicking on the Peer Evaluations link at www.teambasedlearning.org/ndtl.)

Conclusion

Peer assessments and evaluations are essential components of team-based learning. Clearly the process of administering peer evaluation in higher education is a challenging endeavor, especially for instructors new to TBL. As the instructional guidelines and student concerns demonstrate, there is much to consider. As a review of the methods here indicates, there is no single best way to conduct peer evaluation; each method brings with it advantages and potential problems. Nevertheless, there are some basic principles

to keep in mind when establishing a TBL program with a peer evaluation component:

- The skill of performing evaluation is not intuitive. It is useful to assume that most learners have never been taught how to give feedback. Brief written or verbal instruction on how to provide constructive evaluation may prove extremely helpful in allaying students' fears about the process of giving (and receiving) peer review. At a minimum, students should understand the seven characteristics of providing helpful feedback to peers. These characteristics include providing feedback that is descriptive rather than evaluative; being specific, honest and sincere, relevant, timely, in context, desired by the receiver, and concerned with behavior one has the control to change (Michaelsen and Schultheiss, 1988).
- As with any other skill, practice is essential in order to become comfortable with the process. Students need to be able to practice peer review in a safe environment before they can easily apply it for a grade. In most teaching situations, this can be accomplished through a midcourse peer review.
- In many respects, peer review is best received in an environment in which there is a culture of professionalism and a minimal amount of competition and mistrust. The more that courses promote and encourage peer review, the better students will accept it and use it constructively.
- With quantitative evaluations, if students are not forced to discriminate among their teammates, such as giving out only a set number of points, the scores are likely to be highly inflated.
- Students are more comfortable giving qualitative feedback than quantitative feedback. As a result, this might be the easier feedback to begin with for the educator who is reluctant to force a discriminatory quantitative evaluation on the students. However, unless peer evaluations have teeth, groups are vulnerable to students who are prone to social loafing.

Although the process of establishing a peer evaluation system can be frustrating (no one ever got a teaching award for putting together a good peer evaluation), ultimately it is an essential tool for reinforcing the individual accountability so vital to the TBL. Many students need peer review to offset their fear that they will be burdened by having to carry their group. We recommend experimenting with a variety of methods until you find one that works for you in your particular environment. One bit of good news is that if students are accustomed to peer evaluation at an early point in their college careers, the skills will likely transfer across collaborative partnerships in other educational and professional settings (Brooks and Ammons, 2003).

References

Boud, D., Cohen, R., and Sampson, J. "Peer Learning and Assessment." *Assessment and Evaluation in Higher Education,* 1999, *24*(4), 413–426.

Brindley, C., and Scoffield, S. "Peer Assessment in Undergraduate Programs." *Teaching in Higher Education,* 1998, *3*(1), 79–89.

Brooks, C. M., and Ammons, J. L. "Free Riding in Group Projects and the Effects of Timing, Frequency, and Specificity of Criteria in Peer Assessments." *Journal of Education for Business,* 2003, *78*(5), 268–272.

Chen, Y., and Lou, H. "Students' Perceptions of Peer Evaluation: Expectancy Perspective." *Journal of Education for Business,* 2004, *79*(5), 275–282.

Cheng, W., and Warren, M. "Making a Difference: Using Peers to Assess Individual Students' contributions to a Group Project." *Teaching in Higher Education,* 2000, *5*(2), 243–255.

Dochy, F., Segers, M., and Sluijsmans, D. "The Use of Self, Peer, and Co-Assessment in Higher Education: A Review." *Studies in Higher Education,* 1999, *24*(3), 331–350.

Dominick, P. G., Reilly, R. R., and McGourty, J. W. "The Effects of Peer Feedback on Team Member Behavior." *Group and Organization Management,* 1997, *22*(4), 508–519.

Falchikov, N., and Goldfinch, J. "Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks." *Review of Educational Research,* 2000, *70*(3), 287–322.

Gatfield, T. "Examining Student Satisfaction with Group Projects and Peer Assessment." *Assessment and Evaluation in Higher Education,* 1999, *24*(4), 365–377.

Gueldenzoph, L. E., and May, G. L. "Collaborative Peer Evaluation: Best Practices for Group Member Assessments." *Business Communication Quarterly,* 2002, *65*(1), 9–20.

Haberyan, A. "Team-Based Learning in an Industrial/Organizational Psychology Course." *North American Journal of Psychology,* 2007, *9*(1), 143–152.

Lane, D. "Engineering Feedback: A Student-Developed Approach to the Assessment of Peer Evaluation in Civil Engineering." Chicago: National Communication Association, 2007.

Levine, R. E. "Peer Assessment in Team-Based Learning." In L. K. Michaelsen, D. X. Parmalee, K. K. McMahon, and R. E. Levine (eds.), *Team-Based Learning for Health Professions Education*. Sterling, Va.: Stylus, 2008.

Levine, R. E., O'Boyle, M., Haidet, P., Lynn, D. J., Stone, M. M., Wolf, D. V., and Paniagua, F. A. "Transforming a Clinical Clerkship with Team Learning." *Teaching and Learning in Medicine,* 2004, *16*(3), 270–275.

Michaelsen, L. K. "Team Learning: A Comprehensive Approach for Harnessing the Power of Small Groups in Higher Education." *To Improve the Academy,* 1992, *11,* 107–122.

Michaelsen, L. K., Knight, A. B., and Fink, D. L. *Team-Based Learning: A Transformative Use of Small Groups*. Westport, Conn.: Praeger, 2004.

Michaelsen, L. K., and Schultheiss, E. E. "Making Feedback Helpful." *Organizational Behavior Teaching Review,* 1988, *13*(1), 109–113.

Paswan, A. K., and Gollakota, K. "Dimensions of Peer Evaluation, Overall Satisfaction, and Overall Evaluation: An Investigation in a Group Task Environment." *Journal of Education for Business,* 2004, *79*(4), 275–231.

Schelfhout, W., Dochy, F., and Janssens, S. "The Use of Self, Peer, and Teacher Assessment as a Feedback System in a Learning Environment Aimed at Fostering Skills of Cooperation in an Entrepreneurial Context." *Assessment and Evaluation in Higher Education,* 2004, *29*(2), 177–201.

Searby, M., and Ewers, T. "An Evaluation of the Use of Peer Assessment in Higher Education: A Case Study in the School of Music, Kingston University." *Assessment and Evaluation in Higher Education,* 1997, *22*(4), 371–383.

Thackeray, R., and Wheeler, M. L. "Innovations in Social Marketing Education: A Team-Based Learning Approach." *Social Marketing Quarterly,* 2006, *12*(3), 42–48.

Topping, K. "Trends in Peer Learning." *Educational Psychology,* 2005, *25*(6), 631–645.

Topping, K. J., Smith, E. F., Swanson, I., and Elliot, A. "Formative Peer Assessment of Academic Writing between Postgraduate Students." *Assessment & Evaluation in Higher Education*, 2000, *25*(2), 149–170.

CHRISTINA M. CESTONE is from the Department of Educational Psychology at The University of Texas at Austin.

RUTH E. LEVINE is the Clarence Ross Miller Professor of Psychiatry and Internal Medicine in the Department of Psychiatry and Behavioral Sciences at The University of Texas Medical Branch, Galveston.

DEREK R. LANE is an associate dean in the College of Communications and Information Studies and associate professor in the Department of Communication at the University of Kentucky