

Journal of

INDUSTRIAL TECHNOLOGY

Volume 20, Number 1 - November 2003 to January 2004

Using Peer Assessments in Team Activities

By Dr. Steven A. Freeman and Dr. Michael J. Dyrenfurth

Peer-Refereed Article

KEYWORD SEARCH

**Administration
Curriculum
Higher Education
Management
Teaching Methods**



Dr. Steven Freeman is an associate professor in the Department of Industrial Education and Technology at Iowa State University in Ames. He teaches occupational safety courses and coordinates the undergraduate occupational safety option. His research interests are in rural and occupational safety and improving the process of learning and the learning environment.



Dr. Michael Dyrenfurth is Assistant Dean for the School of Technology and professor in the Department of Industrial Technology at Purdue University. His long-term research goal focuses on the refinement of the concept of technology literacy including instrumentation and measurement of the extent to which people possess/manifest it.

Using Peer Assessments in Team Activities

By Dr. Steven A. Freeman and Dr. Michael J. Dyrenfurth

Regardless of the path industrial technology students take upon graduation, they will be required to work effectively as part of a team. To help students develop the skills necessary for this, many industrial technology faculty incorporate team activities into their curricula. Time constraints placed on busy faculty and research supporting the positive benefits of cooperative learning¹ represent two additional reasons for incorporating cooperative group-based learning activities into classes. But, how do faculty assess student progress and assign grades when using team activities? Most who have tried have confronted the inevitable group versus individual problems. While most cooperative learning experts conclude that group activities work best when team grades are adjusted for individual performance (Kaufman, Felder, and Fuller, 2000), this is challenging to accomplish because of the difficulty in ascertaining individual ability and accomplishments within the context of group activities—particularly in larger classes. This article explores the process of using peer assessments in cooperative learning to hold students individually accountable and thereby provide fair grading to students who do the work, as well as to those who do not.

Setting the Stage ***Cooperative Learning***

Johnson, Johnson, and Smith (1998) define cooperative learning as the instructional use of small groups so that students work together to maximize their own and each other's learning. They report research findings indicating that cooperative learning, when compared with competitive and individualistic efforts, typically results in greater efforts to achieve, more positive rela-

tionships among students, and greater psychological health. Referring to cooperative learning as "peer" learning, McKeachie (1999) states that such learning benefits the student both motivationally and cognitively. Motivationally, it has the advantages of interaction with a peer and it provides an opportunity for mutual support and stimulation. Cognitively, it provides students with the opportunity to actively put material in one's own words and begin using the language of the discipline. Haller, Gallagher, Weldon, and Felder (2000), while discussing the advantages of cooperative learning, add that an individual's learning (success or failure) is linked with the learning of the other group members. Felder (1995, p. 32) admits that students new to cooperative learning often "gripe loudly and bitterly about other team members not pulling their weight." However, he also feels that the benefits of group activities far outweigh the problems and that most students will come to accept and enjoy cooperative learning. The authors believe that a key point in making this happen is to address group assessment issues openly and fairly.

Assessment

Huba and Freed (2000, p. 8) state "Assessment is the process of gathering and discussing information from multiple and diverse sources in order to develop a deep understanding of what students know, understand, and can do with their knowledge as a result of their educational experiences; the process culminates when assessment results are used to improve subsequent learning." Traditional assessment techniques often include exams, quizzes, assignments, projects, etc. Additionally faculty often conduct continual informal assessments of their students'

¹ Since, in the authors' experience, most team activities are cooperative/collaborative in nature, the lessons learned from cooperative learning research provide useful guidance.

learning in the classroom. These informal methods include monitoring student questions, comments, body language, and facial expressions (Angelo and Cross, 1993). These authors also describe a variety of techniques to formalize such non-graded classroom assessments.

Regardless of the type of assessment used, experts agree that effective student assessment involves multiple forms of evaluation based on clear educational goals and objectives and where assessment is a learning experience as well as an evaluation device (Huba and Freed, 2000; McKeachie, 1999; Bunta, Lund, Black & Oblander, 1996; Angelo and Cross, 1993). Thus, assessment of group activities needs to include not only assessment of **content** objectives, but also assessment of **group process** objectives. However, a discussion of techniques for the assessment of educational objectives (content or group-based) is beyond the scope of this manuscript. The literature is rich with sources describing generalized assessment processes. The cited sources highlight key points and the authors encourage readers interested in such assessment to refer to these texts for additional detail.

The focus of this article, in contrast, is on how to assess student performance in team, i.e., cooperative group activities. Wankat and Oreovicz (1993, p. 173) state that “Since the group is producing a group report, it is appropriate to give the students a group grade. However, students often feel that is unfair if one student has not done a fair share of the work.” They go on to provide several ways of resolving this problem, most of which require some type of student input. McKeachie (1999) suggests asking each group to submit a single report and then asking each student to rate each group member’s contribution with the understanding that peer ratings may be used to lower the grade of a student whose contribution was perceived to be less than that of the other

group members. Kaufman et al. (2000) implemented a peer evaluation procedure where peer ratings were used to adjust the group grades. Verbal rating criteria were converted to a numerical score. A grade adjustment factor was calculated by dividing an individual’s average peer rating by the team’s average peer rating. The group grade on the project was multiplied by the grade adjustment factor to determine each individual’s grade on the group project. Trytten (2001) used a more detailed peer evaluation process consisting of two sections. First, she asked students to rank group members according to how much they contributed and to identify students who either dominated the activity or relied too heavily on others. Second, she asked eight specific questions that students had to answer concerning each group member. These questions ranged from how often was the group member present at group meetings, to what were their strengths and weaknesses, to a rating (with explanation or justification) of the student’s overall performance. Trytten then created a composite evaluation for each student based on the responses from each group member.

Pragmatic Reasons for Group Activities

From the faculty member perspective, there are practical benefits for using cooperative learning in addition to educational theory-based reasons. Faculty members across the country seem to be faced with increasing demands to do more with less. The requirement to teach more “efficiently” has changed higher education. Out of necessity, many faculty resort to time saving methods of instruction and assessment, such as multiple choice bubble sheet exams. This seems to be particularly true at research universities where the time devoted to instruction must be carefully balanced with the need to publish scholarly work and obtain research funding—both of which are required to survive the tenure and

promotion process. Yet the vast majority of faculty members, regardless of institutional environment, strive to provide the best educational experience possible for their students. Increasingly, this results in the use of contextual-based experiential learning. The most significant benefit of such activities is that students practice their profession in a work-like environment while learning the principles of their undergraduate curriculum. The most significant benefit for faculty is that these learning opportunities, which are often collaborative experiences, provide students with realistic problems in a manner that enables high quality, yet efficient, assessment. In other words, faculty members can do a more thorough assessment of twelve group projects in a given time period than they can assess 60 to 70 individual projects. Yet another significant time savings derives from needing to set up a much smaller number of industry-based experiential learning activities than would be needed for individual assignments. Finally, such learning experiences enhance students’ team skills and provide them with the benefits of cooperative group-based learning.

Case Study

Cooperative learning was implemented as the central methodology in a new 300-level *Safety in Manufacturing* course taught by one of the authors. This is a required course for all industrial technology students at Iowa State University (ISU). It covers broad topic areas in occupational safety and health with a specific emphasis on management responsibilities in a manufacturing setting. The course is a traditional instructor taught course with a web-based supplement². The students in the course are randomly assigned to base groups³ on the first day of class. These base groups were the focal point for all activities outside of the classroom. During the course of the semester, each base group completed four homework assignments, a presentation

²See Freeman and Embleton (2000) for a detailed discussion of the course content and the use of the web-based supplement.

³Base groups as defined by Johnson et al. (1998, p. 6:4) are “long-term, heterogeneous cooperative learning groups with stable membership whose primary responsibilities are to provide support, encouragement, and assistance in completing assignments and hold each other accountable for striving to learn.”

assignment, and a semester-long service learning project as a team. The only activities completed independently by students were daily on-line quizzes and three exams.

Peer evaluations have always been a part of the course but historically they were only done at the end of the semester and, for practical purposes, they only impacted students on the cusp between grades. Individual accountability for group work was based on a variety of informal methods such as non-graded progress reports, interaction with the groups, and on the service learning projects by the group proposals which outlined each member's responsibilities. Groups with significant problems (i.e., students who were not participating) were easily identified. However, there was no formal process of measuring individual accountability or contribution for specific activities. Students were routinely assigned the same grade for group work. Individual adjustments to the group grade were made only in extreme cases, usually impacting students who did not contribute their share to the completion of the activity. Students who contributed more than their share to the activity were seldom recognized with a higher grade. Even under the preceding circumstances, student evaluations of this course have been consistently higher than departmental means. While some students listed "group work" as one of the things they like least about the course, they have been more than counterbalanced by students who listed group activities and projects as what they like best about the course. However, while seldom mentioned in course evaluations, informally some students were expressing concerns about "carrying" group members who were not contributing fairly to group activities.

While the instructor was committed to the principles of cooperative learning, he was never completely comfortable with assigning group grades. He was also not satisfied with the informal processes he was using to adjust group grades for individual performance. For these reasons he started investigating how other faculty

were dealing with this issue. While a variety of methods were uncovered, the processes that seemed most appealing were those involving peer assessments. Based on this action research, criteria for selecting a peer evaluation process were established. The process needed to be: (1) straightforward for the students, (2) not overly time consuming for the students, (3) perceived as fair, and (4) relatively easy to implement.

After teaching the course several times, the *Safety in Manufacturing* course's peer assessment policy was changed to include a specific peer evaluation with each group assignment in addition to the previously used end-of-semester peer evaluation. Each group activity was graded and assigned a "base score" for the group. Individual grades were then adjusted based on the peer evaluations. An average peer rating was calculated for each student as well as an overall average peer rating for the group. Students whose average rating was above the group average received a grade higher than the base group score. Likewise, students with a below average rating received a grade lower than the base group score. Individual grades were assigned to the homework assignments using the following formula:

$$\text{Individual Score} = \frac{\text{Individual Average}}{\text{Group Average}} \times \text{Base Score}$$

Each student's grade on the service learning project was still individually adjusted based on the quality of each student's contribution as defined by the project proposal. The end-of-semester peer evaluation was incorporated into the "class participation" portion of the grading scheme, which accounted for approximately 1/2 of a letter grade. Forms for completing the peer evaluations were adapted from Kaufman et al. (2000). The form used for rating of peer performance for individual assignments is shown in Figure 1. Figure 2 shows the overall end-of-semester peer evaluation form. These forms had the advantage of being easy for the students to complete while using a narrative scale detailed enough

so that students could clearly discriminate between ratings. When scoring, the instructor converted the ratings to a numerical score between 0 (no show) and 10 (excellent).

Results and Discussion

During the semester of the case study, the 41 students enrolled in *Safety in Manufacturing* were organized into eight base groups. Each student had an opportunity to complete five assignment peer evaluations (using the form shown in Figure 1) for the four homework assignments and the presentation assignment. They also completed the overall semester peer evaluation (using the form shown in Figure 2) at the end of the semester. Of the total possible 246 (6 x 41) individual student peer assessments, 234 (95%) were turned in. The remainder of this discussion refers to the analysis of the 234 returned peer assessment surveys. Pearson correlations were used to test for association between average student ratings and student performance in the class. The Pearson correlation coefficient (r) and the level of significance (p) for the correlation were calculated conventionally as described in Snedecor and Cochran (1989).

Average peer ratings (note that peer ratings include self-ratings and ratings from teammates) correlated positively with average exam scores ($r = 0.424$, $p = 0.006$). The correlation between the average rating from teammates (i.e., excluding self-ratings) and average exam scores was slightly stronger ($r = 0.425$, $p = 0.006$) while the correlation between the self-ratings and exam scores was somewhat lower ($r = 0.378$, $p = 0.015$). While the shared variance ranged from only 14% to 18% it should be noted that all correlations were positive. If students read the directions on the peer evaluation forms and correctly applied the ratings, these results would indicate that students who performed best on the exams also

tended to be the most dependable and diligent in contributing to the success of the base group. However, an alternative explanation is that the students who performed best on the exams were rated more positively by their teammates based on ability rather than contribution.

The results and discussion presented below address common concerns for questioning the validity of peer ratings. These include that individuals may inflate self-ratings, that team members may all agree to provide identical ratings, and that personal biases and dislikes may influence ratings (Kaufman et al., 2000).

Self-Ratings Compared to Team Ratings

There was a strong correlation ($r = 0.701$, $p \leq 0.001$) between the average self-rating and the average rating from teammates. By removing four outliers, the correlation between the average self-rating and the average rating from teammates was quite strong ($r = 0.921$, $p \leq 0.001$). A paired t-test showed that over the course of the semester there was no statistically significant difference ($p = 0.042$) between the average self-rating and the average rating from teammates⁴.

Self-ratings were not considered inflated if any other teammate rated that student equally high or higher. Conversely, self-ratings were not considered deflated if any other team member rated that student equally low or lower. Incidences of inflated self-ratings turned out to be less common than the incidence of deflated self-ratings. Out of the total 234 peer assessments returned during the semester, only four times (< 2%) did any student give themselves a self-rating that was higher than the highest rating they received from any of their teammates. These four cases occurred with four different students. Fourteen times during the semester (6%) a student rated themselves lower than the lowest rating they received from any of their teammates. These cases occurred with twelve

Figure 1: Assignment Peer Evaluation Form

ASSIGNMENT PEER RATING OF TEAM MEMBERS

Name _____ Team # _____

Assignment # _____

Please write the names of all of your team members, **INCLUDING YOURSELF**, and rate the degree to which each member fulfilled his/her responsibilities in completing this group homework assignment. The possible **ratings** are as follows:

- | | |
|-----------------------|---|
| Excellent | Went above and beyond—carried more than his/her fair share of the load |
| Very Good | Did what he/she was supposed to do, very well prepared and cooperative |
| Satisfactory | Did what he/she was supposed to do, acceptably prepared and cooperative |
| Ordinary | Did what he/she was supposed to do, minimally prepared and cooperative |
| Deficient | Showed up, but did not help complete assignment, unprepared |
| Unsatisfactory | Failed to show up or complete assignment, unprepared |
| No Show | No participation at all |

*These ratings should reflect each individual's level of participation and effort and sense of responsibility for **this assignment only**, not his or her academic ability or participation in any previous assignments.*

<u>Name of team member</u>	<u>Rating</u>
yourself	self rating

Your signature: _____

Comments:

Adapted from: Kaufman, D. B, Felder, R. M., & Fuller, H. (2000). Accounting for individual effort in cooperative learning teams. *Journal of Engineering Education* 89(2):133-140.

⁴ The four outliers that impacted the normality of the distribution were tested separately using a paired t-test. The t-test also showed no significant difference between the average self-rating and the average rating from teammates for these four students.

different students (two students rated themselves this way twice).

Identical Ratings

Given eight groups each making six peer evaluations, there were 48 opportunities for group members to get together and decide to give everybody identical ratings (i.e., everyone in a team awarding everyone in their team the same rating) and essentially negate any influence of the peer rating process. Teams turned in this type of identical ratings 11 times (23%) during the semester. However, five of the eight teams never turned in identical peer assessments. The occurrence of identical ratings increased as the semester progressed. After the first assignment, Team 5 alone accounted for five of the eleven identical team ratings with their last five peer assessments. The last three peer assessments by Team 6 and the last two by Team 8 were also identical. It is the authors' belief that these identical ratings, particularly as the semester progressed, were indications of effectively functioning teams where the work was being shared equally. It should be noted however, that while the occurrences of identical team ratings were rather rare, individual students gave identical ratings much more frequently. Of the 234 peer assessments returned, 55% of the time students gave teammates (including themselves) identical scores.

Potential Bias

Over the course of the semester, 26 times (11%) a student gave a teammate a rating at least two ratings lower than they were rated by any of the other teammates. These potentially biased ratings occurred to 15 of the 41 students at some point during the semester. Nine students received this type of rating more than once. Of these nine students, four received all of their potentially biased ratings by the same teammate. In one case, a student was rated at least two ratings lower by the same teammate on three of the six peer assessments. This may possibly indicate that the teammate had some personal bias or animosity toward this student. Additionally, there were three occurrences where a student rated a

teammate at least two ratings higher than they were rated by any of the other teammates (including their self-rating). These all occurred with the first peer assessment of the semester.

Semester Ratings Compared to Assignment Ratings

The end-of-semester peer ratings correlated positively with the average of the five assignment peer ratings ($r = 0.579, p d \leq 0.001$). Only six students

Figure 2: Overall End-of-Semester Peer Evaluation Form

PEER RATING OF TEAM MEMBERS

Name _____ Team # _____

Please write the names of all of your team members, **INCLUDING YOURSELF**, and rate the degree to which each member fulfilled his/her responsibilities in completing the group homework assignments and the semester project. The possible ratings are as follows:

Excellent	Consistently went above and beyond tutored teammates, carried more than his/her fair share of the load
Very Good	Consistently did what he/she was supposed to do, very well prepared and cooperative
Satisfactory	Usually did what he/she was supposed to do, acceptably prepared and cooperative
Ordinary	Often did what he/she was supposed to do, minimally prepared and cooperative
Marginal	Sometimes failed to show up or complete assignments, rarely prepared
Deficient	Often failed to show up or complete assignments, unprepared
Unsatisfactory	Consistently failed to show up or complete assignments, unprepared
Superficial	Practically no participation
No Show	No participation at all

These ratings should reflect each individual's level of participation and effort and sense of responsibility, not his or her academic ability.

<u>Name of team member</u>	<u>Rating</u>
_____	_____
yourself	self rating
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

Your signature: _____

Comments:

Adapted from: Kaufman, D. B, Felder, R. M., & Fuller, H. (2000). Accounting for individual effort in cooperative learning teams. *Journal of Engineering Education* 89(2):133-140.

(H²15%) received semester peer ratings that differed by at least one level from the average rating level of their five assignment peer ratings. Four students averaged a higher semester peer rating. Of the two who averaged lower semester ratings, one was a student who, while remaining enrolled in the course, did not participate in the last two homework assignments or the completion of the service learning project. By taking out this one student, the correlation between the semester ratings and the average assignment ratings for the students who completed the course is quite strong ($r = 0.825$, $p \leq 0.001$) and it accounts for 68% of the variance.

Conclusions and Recommendations

This article presented a number of reasons for employing peer assessments. Among these were the need to reduce demands on instructor time, to encourage group participation, and to increase student perception of fairness. Upon reflection, the instructor also noted the importance of developing student comfort with self and peer assessment processes. As industry employs more 360° assessment and other forms of evaluation that include self-assessment, it is important that industrial technology programs inculcate such skills in graduates.

To this end, and to help address the assessment needs of a junior level occupational safety course where all out-of-class activities were completed by cooperative learning teams, a formal peer assessment system was implemented. The rating system fit readily into the structure of the course and was easy for the students to use as a means of holding each group member individually accountable for their contributions to, and participation in, the completion of group activities. The principal observations and conclusions of this case study are presented below:

- Peer ratings exhibited positive correlations with average exam scores. While the collected data does not allow for causal

inferences, the results support the authors' belief that students who show-up prepared and actively participate in class (and consequently do better on in-class exams) demonstrate these same characteristics in their group activities and thus receive higher peer ratings.

- Over the course of the semester there was no significant difference between student self-ratings and ratings from teammates.
- Supporting the results of Kaufman et al. (2000), the results of this case study also suggest that commonly voiced faculty concerns about peer ratings may not be justified.
- Students did not have a tendency to inflate their self-ratings (in fact, deflated self-ratings were more common).
- The majority of the peer ratings turned in by the teams were not identical. And the one team that did consistently turn in identical ratings was a team that appeared to be functioning quite well from the beginning of the semester and consistently turned in the highest quality work.
- The vast majority of the the peer ratings appear to be relatively unbiased and consistent with the ratings of other team members. Only 11% of the ratings were viewed as potentially being negatively biased (i.e., ratings that were at least two ratings below those of other teammates). This is an area that will be investigated more thoroughly in future semesters.

Course evaluations showed an increase, although not statistically significant, in student perceptions of the fairness of the grading system and informal feedback concerning the peer assessment process was quite positive. While the authors are pleased with the results of the peer assessments used in this case study, it is too soon to tell if the assessments implemented impacted student learning, or just made them more comfortable with

the grading mechanism of their group activities. Making the peer evaluation process more effective is an ongoing endeavor. Since completion of the case study peer assessments have been added to each activity associated with the to the service learning project (e.g., proposal, progress reports, final report). Work also continues on exploring ways of incorporating the peer assessments as a means to improve the cooperative learning experience and future research exploring the relationship between peer interactions and learning objectives is being considered. Above all the authors recommend continued exploration of cooperative learning and the means of assessing individual performance in team activities.

References

- Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers* (2nd ed.). San Francisco, CA: Jossey-Bass Publishers.
- Banta, T. W., Lund, J. P., Black, K. E., & Oblander, F. W. (1996). *Assessment in practice: Putting principles to work on college campuses*. San Francisco, CA: Jossey-Bass Publishers.
- Felder, R. M. (1995). Random thoughts...we never said it would be easy. *Chemical Engineering Education* 83(1), 32-33.
- Freeman, S. A., & Embleton, K. M. (2000). Enhancing a residential safety course with the development of an online component: A limited case study. *Journal of Industrial Technology* 16(4), www.nait.org.
- Haller, C. R., Gallagher, V. J., Weldon, T. L., & Felder, R. M. (2000). Dynamics of peer education in cooperative learning groups. *Journal of Engineering Education* 89(3), 285-293.
- Huba, M. E., & Freed, J. E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning*. Boston, MA: Allyn and Bacon.
- Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). *Active learning: Cooperation in the college classroom*. Edina, MN: Interaction Book Company.

Kaufman, D. B., Felder, R. M., & Fuller, H. (2000). Accounting for individual effort in cooperative learning teams. *Journal of Engineering Education* 89(2), 133-140.

McKeachie, W. J. (1999). *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers* (10th ed.). Boston, MA: Houghton Mifflin Company.

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical methods* (8th ed.). Ames, IA: Iowa State University Press.

Trytten, D. A. (2001). Progressing from small group work to cooperative learning: A case study from computer science. *Journal of Engineering Education* 90(1), 85-91.

Wankat, P. C., & Oreovicz, F. S. (1993). *Teaching engineering*. New York, NY: McGraw-Hill, Inc.

