

**Peer Evaluations, and Team Learning In
Undergraduate and Graduate Education**

Bruce A. Rothwell

Abstract

This paper reviewed the available literature concerning the use of student peer evaluations and team learning in undergraduate and graduate education. The literature clearly demonstrates that the simple act of how an instructor assigns groups has a significant impact on the grades their students receive. Also clear from the literature is that training should precede the use of peer evaluations in any classroom setting. Further, student assigned grades become more accurate with training and practice. Although there are several potential problems with the use student peer evaluations and team learning, there are also very pronounced student benefits. Generally, students have been shown to improve their learning, retention, and overall grades when group projects and peer evaluations are correctly used.

Peer Evaluations

"Peer evaluation is the process of having the members of a group judge fellow members on specified traits, behaviors, and achievements" (Sherrard, Raafat, & Weaver, 1994, p. 43). Keith Topping (1998) accomplished an extensive review of literature covering peer assessments at higher education institutions. He reviewed 109 different articles published between 1980 and 1996. Of the 109 articles, 67 were quantitative and presented data the authors reviewed had collected during their research on peer evaluations. The remaining articles that were reviewed were descriptive in nature. His conclusion about the literature prior to 1996 was significant. Although Topping (1998) found a large number of studies had been accomplished, the procedures used and varied disciplines involved in the research made a definitive decision on the effectiveness of peer evaluations impossible.

Even though an overall definitive finding on the value of peer evaluations was not possible, specific aspects of the literature made it possible for Topping (1998) to formulate several meaningful and valid conclusions. The sheer number of different academic disciplines represented by his research lead to the conclusion that peer evaluations may be applicable in all academic areas. Nearly all the studies reviewed indicated that the instructors had

mandated the peer evaluations instead of the students initiating them. Although the process where instructors who pushed the peer evaluation system on their students was seen as a possible problem, the process of involving students in peer evaluations resulted in students getting more involved in the subject matter. Peer evaluations created a sense of student ownership and required the students to take responsibility for the evaluations they provided (Topping, 1998).

Although a definition of peer evaluations is easily definable and generally agreeable to most, implementation strategies and the value placed on them differs greatly. Some instructors see them as merely beauty contests while others have found them to be an important part of the learning process. In part, those who have used and liked peer evaluations found that they lent a unique perspective that the instructor could not possess. Specifically, peer evaluations come from those individuals closest to the individual being rated and who are also the people who have the most contact with them (Sherrard, Raafat, & Weaver, 1994).

When instructors assign people to team projects, the professor must understand that there will invariably be some students who will attempt to ride the coattails of the rest of the group. These free-riding individuals do not participate as

much as the remainder of the group and in some cases may not participate at all (Dyrud, 2001). Left alone, student groups may or may not effectively solve the various types of problems they can encounter. In many cases, the instructor is not even aware of discontent within the student teams until very near the time for the project to be delivered for a grade. Generally, near the end of term, the students are frustrated and group interaction may be well beyond the point of simple disagreement. One way to alleviate some of the problems found within group assignments is for the instructor to use a system of peer reviews within the group (Dyrud, 2001).

Terry Gatfield (1999) suggests that groups permeate the working environment because a group's many different talents allow it to accomplish what no individual person could do alone. A natural outgrowth of the group-working environment was to bring the use of groups into the classroom where they can improve student learning. The use of groups in a class is reported to have many advantages including the finding that students will learn first-hand how to interact within a group and become better prepared to take their place in a work environment where groups have become prevalent. Following logically to the next level, the use of groups in the class created a need for instructors to understand the contribution of each individual student. To solve this need, a system of peer

assessments was established and used in various classrooms.

Most instructors found that creating teams in the classroom was the easy part. However, once the groups were created, one of the main problems then facing the instructor was how to grade the efforts of a group project. The final grade for an individual project is a relatively easy process of reviewing the work and assigning an appropriate grade. Because only one person worked on the assignment, the grade received should be a fair estimate of the effort the student put into the assignment. However, when more than one person has contributed to the completion of a project, there are little formal means for the instructor to know how much of a contribution was made by each individual. Therefore, one or more of the students may receive a free ride while others in the group did more than their fair share. Conversely, it is possible that a poor final project may have been submitted from a team comprised of several members who did little and one student who tried exceptionally hard. Any time an instructor assigns the same grade to all members of a team, it is probable he or she will penalize some and reward others based on factors outside of the students' control (Maranto & Gresham (1998).

Many studies have found significant benefits derived from using peer evaluations. If used correctly, peer evaluations will have a positive impact. It is possible for the peer evaluation process to

improve student grades and their overall involvement in course activities (Topping, 1998). Additionally, when students know that their efforts are going to be judged by fellow classmates, who may also be their friends, the students may be motivated to try harder than if they are only being evaluated by the instructor (Hite, 1996).

One important way of assisting the instructor in the assignment of a fair grade for group assignments is to have the members of the team tell the instructor how much of an effort each member of the team made to the project. The process of students reporting on other students is referred to as peer evaluations. Sherrard, Raafat, and Weaver (1994) performed an analysis of students who received peer evaluations and an instructor assessment for the same in-class presentation. The peer evaluations received determined 20 percent of the students' final grade in the course. In addition to the peer evaluations, each student rated their own performance on two in-class presentations. The researchers found that the self-evaluations the students provided were very accurate in comparison to evaluation scores they received from their fellow students. The researchers also found that only gender had an impact on the evaluations students received with females rating individuals higher than males. Finally, each student accomplished an end of course critique where they were asked specific questions about the peer evaluation process. The

majority of students reported that the peer evaluations were a valuable means of providing feedback and indicated that they had actually learned from the evaluations. However, the majority of students also felt that their peers had evaluated them on factors other than just the presentation on which they should have been assessed (Sherrard, Raafat, & Weaver, 1994).

In a different study, a researcher had students predict how fellow students would do on a final examination based on their perceptions of them during the term. Students had only in-class discussions and seminar presentations on which to base their predictions. Although the study only involved 75 undergraduates in their third year, the results reported were noteworthy. The researcher found that the students' peers accurately predicted how the students would perform on the final examination. The results were even more accurate of a predictor when the one being evaluated and the one doing the evaluation were most similar in ability and performance during the class. In other words, the closer the rater and the one rated were in the final class ranking, the closer the prediction was to where they would finish (Orpen, 1994).

In a study of 59 graduate students, Kelmar (1993) found that the students fairly accurately assessed the performance of their peers on an in-class presentation of an outside reading assignment. Unlike many studies, this research used the results of the peer assigned evaluations to determine 15

percent of the students' final grade. The students were not aware that their instructor would also be grading the presentation and would compare the student provided evaluations to his. Before the assessment, the students were trained in the dynamics of peer evaluations. Kelmar (1993) found that the student assessments were on average, significantly higher than his were. Further, the instructor found six of the presentations to be unacceptable graduate level work while none of the students assigned a failing grade to any of their peers. Although the students' evaluations were higher on average, they were all tightly centered about the mean, whereas the instructor's ratings had a much larger variation. The author attributed the differences in scoring to three main reasons. The students were seen as more sympathetic to their peers, the students were not previously experienced in assigning grades, and the students did not have the same opportunities to converse with other raters as the instructor had to discuss with his peers (Kelmar, 1993).

Marilyn Dyrud (2001) used a series of peer reviews designed to improve group interaction and report the contribution level of each member within the group. To this end, she used three mandatory peer reviews. The first two were open evaluations and were designed to let the group solve their own problems, promote group interaction, and eliminate undesirable group behavior as early as possible in the group project. The final peer review was

only seen by the instructor and was the one that she used to assign a portion of the final grade to the various group members. Group work was seen as essential by the researcher to prepare her students for what they will experience in the majority of work settings (Dyrud, 2001).

Gatfield (1999) studied 261 undergraduate students and found that the students were very satisfied with the system of peer evaluations used in his class. Further, he found that students who had previously been in an actual working environment had a higher level of group satisfaction than those students who had no previous work experience did. Lastly, he found there was no statistical difference between males and females in acceptance of groups and peer evaluations in the classroom. The one caution that the author presented was that groups and peer evaluations might not be appropriate for first year undergraduates because they may not possess the necessary prerequisites to handle group dynamics.

Persons (1998) determined whether or not factors that existed prior to the start of her accounting classes and those factors acquired during the accounting class had an impact on the peer evaluation students received. She called the factors that existed prior to being placed in a group learning setting ex-ante factors and those that were acquired while in the group ex-post factors. The ex-ante factors included gender, race, GPA and prior accounting background. The

ex-post factors included test scores, participation in class, and grades for group assignments. The researcher, who tried to distribute the various ex-ante factors as equally as possible, assigned the groups. All peer evaluations were conducted at the end of the course. Persons (1998) found that gender, race, and previous accounting knowledge had no impact on the peer evaluations received. However, the students' previous GPA and declaration of an accounting major were positively related to the peer evaluations they earned. When looking at the factors that students acquired since the start of the course, only their participation and group homework grades were positively related to their peer evaluation scores. The major findings from this study were that future instructor assigned learning teams should be equally filled based on previous GPAs and declaration of a relevant major (Persons, 1998).

Many instructors have developed creative ways around the problem of assigning group and individual grades for a group assignment. Most of these means center around some form of peer evaluations. One such peer evaluation system is known as the Knickrehm Method. Within the Knickrehm Method, the instructor assigns a group grade and members of the group assign a specific number of points to all other members of the group based on the students' contribution to the group project. Each member of the group is allocated a

specific number of points that they award to anyone in their group other than themselves (Maranto & Gresham, 1998).

Typically, the Knickrehm style has point ranges from zero points, awarded to members who did not contribute, to a maximum of four points, for members who did most of the work. Members are allocated a specific number of points sufficient to award everyone on their team with two points and still have one point left over to award to the best performer. In this manner, the majority of the members are awarded two points and one person receives three. The description for award of two points is that the person contributed their fair share. Although the awarding of two points is the norm, each person can give any other member zero or one point, leaving a higher possible score for someone else (or more than one other person) of their choosing. Because the instructor limits the percentage of the final grade that the peer evaluations impact, the points assigned by peers within the groups can only change an individual's grade on the margins and generally never more than ten percent (Maranto & Gresham, 1998).

Maranto and Gresham (1998) reviewed teams at two different universities to determine the impact of using the Knickrehm Method for peer evaluations. The first author found high student satisfaction with the method and in over 40 groups comprised of between four and seven members had only one student

complaint concerning the evaluations. Findings from the first author show that those scores that were reduced were lowered from four to eight percent and scores that were increased were improved from eight to 16 percent. Seldom did these changes have an impact on the final grade assigned. The second author was working with much larger class sizes and was met with high levels of student dissatisfaction. Although there were more students' complaints, about ten percent of the grades were lowered and 38 percent were increased through the peer evaluations. Part of the reason for the higher student dissatisfaction in the second scenario was that the instructor was at a university using the plus-minus grading system. Therefore, any movement in these peer evaluations away from the Knickrehm normal score of two had an increased potential for changing the students' final grade. The overall conclusion from this study was there might be more appropriate settings than others for using peer evaluations (Maranto & Gresham, 1998).

Peer evaluations can be effectively used on several different assignment types including oral presentations, group projects, and individual writing assignments. "An important feature of most WAC [Writing Across the Curriculum] programs is the use of peer review and peer grading; that is, students' evaluation of the writing efforts of their peers" (Kerr, Park, & Domazlicky, 1995, p. 357). One research study reviewed

by Kerr, Park, and Domazlicky (1995) found that when undergraduate business communication instructors used peer evaluations of writing assignments, the majority (82 percent) had their students review peer papers prior to the student turning in the paper for a grade. The remaining 18 percent had fellow students assign some form of an assessment of the paper turned-in that was used in the determination of a final grade. There are several positive reasons for having each student review their peers' papers. One of the main advantages is that by reading other papers, the students are exposing themselves to additional information concerning the course topic. The additional exposure to course materials helps retention and understanding of key points. Further, by knowing that their peers will see their work, the student will have an increased incentive to do their best because they do not want to look bad in front of their friends and peers. Closely related to this benefit, the students, because of the closer relationship they have with their peers, might place more weight on the opinions of fellow classmates than in their instructor's evaluation.

Although most studies involving peer evaluations involve performance within a group setting, one study tracked the evaluations of sophomore and junior undergraduates' evaluations of two individual writing assignments accomplished over the course of a single term. Although

the results showed that the students' evaluations consistently exceeded that of the instructor's evaluations; the differences were most pronounced in the scoring of the first paper. By the second paper of the term, peer evaluations decreased while the instructors' evaluations had increased. The increased instructor evaluations were seen as caused by the students' improved writing abilities resulting from the accomplishment of the first paper and the additional learning that took place from the review of their peers' first paper. One of the main findings of this study was that by learning how to evaluate others, the students also learned how to more critically evaluate their own efforts. The decrease in the students' evaluations of their peers was seen as being caused by learning what to look for on the first paper. Therefore, the students' ability to evaluate improved with practice (Kerr, Park, & Domazlicky, 1995).

Potential Problems with Peer Evaluations

Peer evaluations generally create significant initial anxiety in both the instructor and the student (Topping, 1998). Therefore, before attempting to implement a team learning approach, the professor must understand that there is much more to the process than simply assigning students a group project. The instructor has to provide the foundation and continuously follow-up to make sure the team learning approach is successful (Hite, 1996).

In recent years there has been an increasing trend to involve students in the evaluation process. Generally, these models fall into the two general areas of either self-evaluation or peer evaluation. Two of the biggest problems with the growing use of peer evaluations are that the students doing the evaluations may not be qualified raters and also may not be impartial in assigning their evaluations. Simply stated, some students may rate others of their same sex or race differently than those of the other sex or races (Ghorpade & Lackritz, 2001).

Ghorpade and Lackritz (2001) reviewed the performance of 221 senior level undergraduate students in human resource management to determine if there were differences in the way fellow students rated their peers based on sex and race when compared to previous class participation. The assignment that was peer evaluated was an in-class presentation provided by a team of three to five students. Each student received a peer evaluation on both an individual and team basis. The authors also evaluated each student. They then compared the amount of student participation they observed against their evaluations and against the evaluations provided by the students based on sex and race differences. Ghorpade and Lackritz (2001) also had each student self-evaluate their participation. The researchers then compared the students' self-evaluation with the authors' assessments of the student. In

85 percent of the cases, the students' self-evaluation matched that of the researchers. The researchers' findings also clearly showed that the more students participated in their classes, the higher that students' peer evaluations were. In fact, "[t]he single most significant influence behind the ratings process was frequency of participation in classroom discussions by the presenters" (Ghorpade & Lackritz, 2001, p. 279). Also interesting, was that there was no difference in the rating received or given based on the sex of the individual. In other words, men and women rated each other the same. Although there were no differences found for gender, the same was not true for differences in races, with African Americans receiving the lowest scores. Further, the Asian Americans did significantly better than any other group. Strangely enough, the highest ratings for whites came from the Asian and African American groups. Although not an initial consideration of their study, Ghorpade and Lackritz (2001) found that age also had an impact on the ratings students received. The older the student doing the presentation, the higher the rating they received. The overall finding of their research was that peer evaluations should not be used as the sole means of grade assignment.

Lejk, Wyvill, and Farrow (1999) report that a vast majority of United Kingdom professors (95 percent) reported using some form of group assessment at least once. Further, they indicated that the majority also

felt an "...uncertainty about the reliability of group assessment, especially when all group members are allocated the same grade" (p. 5). What makes the Lejk, Wyvill, and Farrow (1999) study significant was that their research spanned four years and tracked 729 university students who they had assigned to work within groups. About half of the 729 students were assigned to groups based on their performance on two tests they took before being assigned into their groups. The researchers separated the test scores into three classifications; low, medium, and high and assigned students into their groups using these three classifications. One half of their students were teamed with other students within their same classification. Specifically, students with high scores were teamed with other students with high scores while students with low scores were assigned to teams with other low scoring students. The other half of their students were assigned within groups of mixed performance results on the first two tests. These teams were comprised of three students with one from each of the high, medium, and low performance categories.

The results the researchers found were striking. Students who had done the best on the first two tests averaged 11 percent lower marks if they were assigned to mixed groups than those who had previously done well and were assigned with others who had also done well. Students who had done poorly on the first two tests scored an average of 12 percent higher when they

were assigned to mixed groups than those who were assigned with others who had not initially done well. The implications of these findings are clear, "[t]he method by which a group is formed seems to have an effect upon the performance of the group" (Lejk, Wyvill, & Farrow, 1999, p. 13). The act of simply assigning students to groups has a tremendous impact on the students' final grade and more importantly the amount of learning that they take away from a class. Randomly assigning students to groups would tend to push all scores toward the middle. While assigning by ability will hurt either the poor or high performers depending on whether they were assigned to mixed groups or groups of equal abilities.

Cheng and Warren (1999) determined if there were differences in the scores provided by students and instructors for first year electrical engineering students. The researchers first trained their students on what they should look for when evaluating their peers. Separate assessments were provided from both the instructors and the students for each written and oral assignment. The researchers found a significant difference between the grades provided by the instructors and the students. The students consistently grouped scores together with little variation in grades. The instructor assigned grades had a greater variation with a larger range of scores. Although there was an initial significant disagreement between the student and instructor assigned grades, as

the students graded more assignments, the differences started to shrink. The authors saw the narrowing of the differences between instructor and student assigned grades as occurring naturally as the students gained experience in grading (Cheng & Warren, 1999).

Cheng and Warren (1999) were not the only authors who found that students tend to assign grades for their peers that were clustered around the mean with very little variation. Goldfinch, Laybourn, MacLeod and Stewart (1999) also found very limited variation in scores provided on one's peers within their teams. Topping (1998) reviewed 25 articles that compared student provided peer evaluations to ratings provided by their instructors. He found that many studies reported that their peer evaluations clustered around the median. Therefore these researchers were in agreement that student evaluations pushed the poorest and highest performers toward the middle.

Students may not accept the concept of other students rating them. Poor group performers may not believe the evaluations they receive from fellow students (Topping, 1998). The conclusion here was that any time an instructor requires a peer evaluation, the instructor must stay involved in the process. Initial instructions followed by continuous monitoring of the situation is required to uncover and fix problems as soon as possible (Topping, 1998).

The grade a student normally receives on their individual projects may differ significantly from the grades they receive on a group project. There are several forms of traditional grading including tests and group projects. A relatively new means of grade determination used in some business schools is the assessment center. The idea behind assessment centers is the use of different exercises that are a reflection of what a manager might experience during his or her normal business day. A major flaw in using traditional forms of assigning grades is that they only determine the declarative and the knowledge compilation earliest stages of learning. Where the assessment center captures all stages of learning including the elusive proceduralization stage that allows application (Bartels, Bommer, & Rubin, 2000). The researchers determined whether or not there was a relationship between 347 undergraduate students' GPAs and how they did on an assessment center. The researchers make it clear that an individual's GPA is impacted by many different things such as motivation and interest instead of simply a matter of intelligence. However, they found that scores on the assessment center correlated with most other scores provided during the class including GPA, tests, and discussion. The only scores that did not correlate with the assessment center score were the scores they received for the group projects. The reason provided for a lack of correlation on the group project was

the fact that the group grade was not done on an individual basis where the assessment center grade was always individually graded (Bartels, Bommer, & Rubin, 2000).

Cooperative and Group Tests

Some research studies show that instructors can further improve the quality of student team learning through the use of group tests. Guest and Murphy (2000) state that prior "...research on the nature of memory suggests that traditional written individual examinations may not maximize long-term retention of information and concepts" (p. 350). They studied 90 graduate students in a teaching program. The researchers developed a group verbal final exam that determined mastery of course materials. They collectively designed the group test so that it required student application of key points from throughout the term. The researchers compared the results from the group exams to students who took the test individually. During the group exam, any student could be called on to respond to any question and their response would be the only grade their group received for that particular test question. At the end of the exam, all members of the group provided critiques of the test format. From these critiques, less than ten percent of the students had any negative comments and the negative comments were generally not related to the group nature of the test. Rather, the

criticisms were centered on the additional pressures the students felt in responding for the entire group. The vast majority of students reported that they felt they would retain more of the information from this testing format than they would from other means of learning (Springer, Stanne, & Donovan, 1999).

Specific findings from the Springer, Stanne, and Donovan (1999) study were that students who were in the group test classes reported spending significantly more time preparing for the final exam than students who were in an individual setting. The authors felt this was a result of the interdependence the students felt and not wanting to let down their team members. Further, the authors felt the extra time spent studying aided retention and contributed to the students reporting that they felt they would retain the information more than in other classes. Finally, the researchers indicated that although there were significant benefits to a group test, professors should not use the group exam as a significant means of assigning student grades. Instead, the group test should simply be one of several different measurements.

Hite (1996) conducted a study of 278 undergraduates by separating them into fairly equally divided control and experimental groups. The control group took three course tests and one final examination as individuals. The experimental group accomplished the exact same tests as the individuals accomplished.

The difference the experimental group received was that they also accomplished the same three course tests in a group setting on the day after they took the tests as individuals. Instead of repeating the exams, the control group spent the entire next in-class session reviewing the tests. Therefore, the amount of in-class time spent on the tests was about equal. The experimental group took the final exam on an individual basis and did not have the same test review as the control group. The experimental groups were comprised of three students with one high, one middle, and one low performing individual in each group. Composition of the groups changed after each test based on the scores from the previous test(s). All students took this final on an individual basis. Although the final exam was comprehensive, no test questions were repeated for either the control or experimental groups. Results clearly showed that the experimental group did significantly better on the final exam than the control group. The main conclusion reached in this study was that the students learned more and had better retention from the group nature of the previous tests. In addition to differences in final exam grades, Hite (1996) found that the end of term student critiques from the experimental group was more favorable than from the control group.

Team Learning

“What students learn is greatly influenced by how they learn, and many

students learn best through active, collaborative, small-group work inside and outside the classroom” (Springer, Stanne, & Donovan, 1999, p. 21). Learning within teams is effective because small numbers of students work together and help one another succeed. It is the significant positive influence brought to bear by the students’ peer group in an academic setting that provides the foundation for improved learning and retention. Cooperative learning is more effective than traditional means of teaching, in part, because team learning motivates and actively involves the individual in their own education. In addition, a group learning environment provides the student with the opportunity to see how their peers handle the same situations they are exposed to. Seeing how others problem solve gives the student the chance to adjust their own techniques and modify their views when presented with conflicting view points (Hite, 1996).

Springer, Stanne, and Donovan (1999) found that students derived several significant benefits when they were taught in a group setting. Not the least of these findings was that undergraduate science, mathematics, engineering, and technology students had a 22 percent higher rate of continuing in their major if they were taught in a group. Such an increase is important because in the science, math, engineering, and technology fields, there is an exceptionally high rate of students who change majors.

Springer, Stanne, and Donovan (1999) also found that undergraduates who were in a small group setting did statistically significantly better than individuals who were not in a group learning environment. Further, these researchers found that students from the group learning classes reported more favorable feelings toward the subject matter they were exposed to. In part, the reasoning behind their findings was that the students may place more value on the group succeeding, may support one another more, and would learn more from one another in a group setting than individually (Springer, Stanne, & Donovan, 1999).

There were no differences in the successes of male and female students within group learning courses. Further, African American and Latino students did better in undergraduate science, mathematics, engineering, and technology classes than those same minorities who were not in a group learning environment (Springer, Stanne, & Donovan, 1999). Their study also found no difference between cooperative and collaborative group teaching styles. The authors defined the cooperative approach as being more instructor structured than the collaborative approach, which relies more on the groups to determine how they will accomplish the assigned tasks. The important thing was not the type of teaching style that was used for the group but rather that there was a group (Springer, Stanne, & Donovan, 1999).

When students are assigned to teams without prior team training, they may not learn as much as if they were first taught about group learning and then allowed to work within a team. Further, many instructors are not totally sure how to properly assess the accomplishments of a team or the individuals that make up the team. In a 1999 study, Goldfinch, Laybourn, MacLeod and Stewart addressed all of these problems. The authors first taught the students what teams were and what makes them function effectively. Next, they brought in several local employers and trained them in what they should be looking for during future periods of student assessment from an individual and team perspective. Lastly, the authors used the employers to provide feedback to their students while the students were still at an early stage in their group projects. The perspective of the employers was found to be important to the students because they had an air of realism. In addition, they had gained the immediate respect of the students because they were already assessing employees who were working in teams. The major findings from this study were that the students seemed to more readily accept the team concept after they received training and feedback from the practitioners. Further, when the same students were seen in later classes, they were actually using the principles they learned about working within groups in the earlier classes (Goldfinch, Laybourn, MacLeod & Stewart, 1999).

The majority of students reported improved learning when they worked within groups. In a study where 140 undergraduate computer science students were equally divided between group and individual learning situations, those assigned to the groups reported that they felt they had learned more. Although the students in the study reported that they had learned more, their grade distribution showed there was no difference between group and individual learning (Benbunan-Fich, 1999).

References

- Bartels, L. K., Bommer, W. H., & Rubin, R. S. (2000). Student performance: Assessment centers versus traditional classroom evaluation. *Journal of Education for Business*, 75, 198-201.
- Benbunan-Fich, R. (1999). Assessing learning of asynchronous computer-mediated communication in the classroom. *The Journal of Computer Information Systems*, 39, 82-87.
- Cheng, W., & Warren, M. (1999). Peer and teacher assessment of the oral and written tasks of a group project. *Assessment and Evaluation in Higher Education*, 24, 301-314.
- Dyrud, M. A. (2001). Group projects and peer review. *Business Communication Quarterly*, 64, 106-112.
- Gatfield, T. (1999). Examining student satisfaction with group projects and peer assessment. *Assessment and Evaluation in Higher Education*, 24, 365-377.
- Ghorpade, J., & Lackritz, J. R. (2001). Peer evaluation in the classroom: A check for sex and race/ethnicity effects. *Journal of Education for Business*, 76, 274-281.
- Goldfinch, J., Laybourn, P., MacLeod, L., & Stewart, S. (1999). Improving groupworking skills in undergraduates through employer involvement. *Assessment and Evaluation in Higher Education*, 24, 41-51.
- Kelmar, J. H. (1993). Peer assessment in Graduate Management Education. *The International Journal of Education Management*, 7, 4-6.
- Kerr, P. M., Park, K. H., & Domazlicky, B. R. (1995). Peer grading of essays in a principles of microeconomics course. *Journal of Education for Business*, 70, 357-363.
- Hite, P. A. (1996). An experimental study of the effectiveness of group exams in an individual income tax class. *Issues in Accounting Education*, 11, 61-59.
- Maranto, R., & Gresham, A. (1998). Using "World Series shares" to fight free riding in group projects. *PS, Political Science & Politics*, 31, 789-791.
- Persons, O. S. (1998). Factors influencing students' peer evaluation in cooperative learning. *Journal of Education for Business*, 73, 225-229.
- Orpen, C. (1994). Perceived similarity: Its effect on the accuracy of peer evaluations among university students. *The International Journal of Education Management*, 8, 4-6.
- Guest, K. E., & Murphy, D. S. (2000). In support of memory retention: A cooperative oral final exam. *Education*, 121, 350-354.

- Lejk, M., Wyvill, M., & Farrow, S. (1999). Group assessment in systems analysis and design: A comparison of the performance of streamed and mixed-ability groups. *Assessment and Evaluation in Higher Education*, 24, 5-14.
- Sherrard, W. R., Raafat, F., & Weaver, R. R. An empirical study of peer bias in evaluations: Students rating students. *Journal of Education for Business*, 70, 43-47.
- Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*, 69, 21-51.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68, 249-276.