

7-2011

Impact of Peer Evaluation Confidentiality on Student Marks

Christina Hamme Peterson
Rider University, cpeterson@rider.edu

N.Andrew Peterson
Rutgers, The State University of New Jersey, andrew.peterson@ssw.rutgers.edu

Recommended Citation

Peterson, Christina Hamme and Peterson, N.Andrew (2011) "Impact of Peer Evaluation Confidentiality on Student Marks,"
International Journal for the Scholarship of Teaching and Learning: Vol. 5: No. 2, Article 13.
Available at: <https://doi.org/10.20429/ijstl.2011.050213>

Impact of Peer Evaluation Confidentiality on Student Marks

Abstract

Although group work is commonly used in university-level instruction, social loafing, domineering team members, poor attenders, and inequitable distribution of marks have been identified as obstacles to team-based learning. Peer evaluation has been proposed as one vehicle to address these issues. For use in grading, peer evaluations are often anonymous; however, as tools to address team functioning, they should not be conducted anonymously, but rather with the results discussed as feedback. It has been suggested, however, that non-confidential peer evaluations will artificially elevate students' marks. In this study, we investigated the impact of peer evaluation confidentiality on students' marks. Without a weighted correction, confidential evaluations significantly dropped students' marks while non-confidential evaluations raised them. Implications for practice are discussed.

Keywords

Peer evaluation, Marks, Confidential, Social loafing, Team, Group

Impact of Peer Evaluation Confidentiality on Student Marks

Christina Hamme Peterson
Rider University Lawrenceville,
New Jersey, USA
cpeterson@rider.edu

N. Andrew Peterson
Rutgers University
New Brunswick, New Jersey, USA
andrew.peterson@ssw.rutgers.edu

Abstract

Although group work is commonly used in university-level instruction, social loafing, domineering team members, poor attenders, and inequitable distribution of marks have been identified as obstacles to team-based learning. Peer evaluation has been proposed as one vehicle to address these issues. For use in grading, peer evaluations are often anonymous; however, as tools to address team functioning, they should not be conducted anonymously, but rather with the results discussed as feedback. It has been suggested, however, that non-confidential peer evaluations will artificially elevate students' marks. In this study, we investigated the impact of peer evaluation confidentiality on students' marks. Without a weighted correction, confidential evaluations significantly dropped students' marks while non-confidential evaluations raised them. Implications for practice are discussed.

Keywords: peer evaluation, marks, confidential, social loafing, team, group

Introduction

Group-based learning is common practice in university classrooms. Research suggests that, given the appropriate circumstances, such group work yields better student performance, lower student anxiety about difficult content, higher self-efficacy (O'Brien, 1995; Slavin, 1995), and greater student learning (Carlsmith & Cooper, 2002; Longmore, Dunn, & Jarboe, 1996; Revere, Elden & Bartsch, 2008). In addition, for those who prefer a more interactive work environment, group-based learning increases student engagement and perceptions of task importance (Peterson & Miller, 2004). Despite its ubiquity, the prospect of group work is often received by students with feelings of ambivalence if not outright hostility, in large part due to actual or anticipated interpersonal conflict and negative outcomes (Colbeck, Campbell, & Bjorklund, 2000; Monk-Turner & Payne, 2005). Two objections to group work have been well documented in the literature. These are concerns about social loafing (e.g. Ashraf, 2004; Burdett, 2007; Carlsmith and Cooper, 2002; Myers, Smith et al, 2009; Meyers, 1997) and unfair grading procedures (e.g. Hoffman & Rogelberg, 2001; Johnson & Johnson, 1992; Johnson, Johnson & Smith, 1998). Also reported by students as critical to the quality of the student group experience has been the presence of domineering teammates (e.g. Burdett, 2007; Jalagas & Sutton, 1984), and poor attendance and/or preparation of teammates (e.g. Burdett, 2007; Feitchner & Davis, 1984).

Problems in Group Work

“Social loafing” is a phrase coined by Latane, Williams, and Harkins (1979) to describe the phenomenon of a decrease in individual effort due to the presence of others. In group work, individual outputs are not clearly identifiable, thereby providing a screen for poor performance and a lack of incentive for performance excellence (Harkins & Jackson, 1985). As a result, some students will make the decision to allocate their time and effort to individual tasks, where their contribution will be identified and evaluated, and to contribute minimally to the group so as to reap the benefits with a modicum of effort. The elements of the group task consequently are inequitably distributed among group members, with some members shouldering substantially more of the burden (Ashraf, 2004; Carlsmith & Cooper, 2002; Myers, et al., 2009). Despite the prevalence of social loafing, students report that they rarely confront the loafers (Myers, et al, 2009), and feel at a loss as to how to handle such dynamics (Burdett, 2007).

Concern about the nondiscrimination of marks when contribution to student assignments has been inequitable is widespread (Ashraf, 2004; Hoffman & Rogelberg, 2001; Johnson, Johnson, & Smith, 1998). Research suggests that equal marks for all group members can lead to higher marks for poorer students and lower marks for stronger students, resulting in more resistance to group work from high-performing students (Ashraf, 2004; Hoffman & Rogelberg, 2001; Kagan, 1995) and more favorable attitudes from low performing students (Su, 2007). Burdett (2007) found that shared group marks were often challenged by students who felt that such a marking method was unfair and Johnson and Johnson (2008) indicate that the failure to provide individual-level marks in group projects increases social loafing.

Social loafing is discussed in the literature as stemming from the invisibility of individual contribution to the final product and consequently the decision to reduce effort on the basis of a cost/benefit analysis. The role of the domineering student identified in the research, however, may provide another possible explanation for this phenomenon. The domineering student is one who has a clear and inflexible plan for how the task should be completed. He or she is quite vocal in directing group members and behaves as if believing that the only way to ensure a quality output is to dictate all aspects of the task or to complete them him or herself (Burdett, 2007). The perspectives of other team members are largely ignored, team member input is not solicited and the domineering student is fairly rigid in his or her ideas about how to complete the task. In the worst case scenario, students faced with such a teammate are left feeling incompetent and incapable and may revert to social loafing merely in response (Jalajas & Sutton, 1984).

Finally, students report another difficulty in group work as stemming from the failure of individuals to attend meetings or to be prepared when they do attend (Burdett, 2007; Feitchner & Davis, 1984). This phenomenon may occur irrespective of the social loafing and/or domineering role and presents a challenge to completing the required assignments.

The Use of Peer Evaluation to Address Group-Work Problems

Peer evaluation and the incorporation of peer based scores into the final mark for each individual has been proposed to address issues of social loafing and inequitable mark distribution (Asgari & Dall’Alba, 2011; Cheng & Warren, 2000; Kaufmann, Felder & Fuller, 2000). Such evaluations have been found to have validity due to high correlations with teacher ratings (Falchikov & Goldfinch, 2000) and supervisor ratings (Conway and Huffcutt,

1997) and clear differentiation of ratings based on group member's contribution (Saito & Fujita, 2009). There are two main types of peer evaluations in the literature: those that rank team members and those that rate team members. In the ranking method, group members rank their teammates from best to worst (Bushell, 2006; Pope, 2001) or divide an even number of points amongst their teammates, with those contributing more awarded more points and those contributing less receiving fewer (Carson & Glaser, 2010; Longmore, Dunn, & Jarboe, 1996; Maranto & Gresham, 1998). In the rating method, each team member is rated on a set of performance characteristics and these ratings are independent of the ratings of fellow team members (Li, 2001).

It has been suggested that peer ranking methods more clearly differentiate high and low performers than do peer rating methods and so are recommended for use in the grading process (Falchikov & Goldfinch, 2000; Lejk & Wyvill, 2001). If differentiation of students on the basis of contribution and the consequent elimination of social loafing is the sole purpose of peer evaluation, then peer ranking as a unidimensional measure of student performance (typically, the level of contribution) does indeed seem to meet that end. However, process characteristics of a team are strong predictors of successful performance (Campion, Papper & Medsker, 1996) and, therefore, the improvement of interpersonal communication, cooperation and teamwork skills is often cited as a goal of group-based learning (Asgari & Dall'Alba, 2011; Bacon, Stewart, & Silver, 1999). It is these skills that enable students to begin to address issues not only of social loafing, but also of domineering members and lack of attendance or preparation. If we argue that the objective of team-based work in education is to prepare students for the work environment, peer evaluation rankings establish an artificial method for managing the social loafing group dynamic. They eliminate the social loafer without intervention on the part of group members, do not facilitate the learning of teamwork skills and may serve to prevent the circumstances and discussion necessary for group members to actively address dynamics that hinder performance.

Boud, Cohen, and Sampson (1999) argue that assessments used in a course must be aligned with the kinds of learning the course aims to promote. In that case, it is critical for faculty to recognize that the level of an individual's contribution is not likely a learning objective of a group project but is rather a symptom of group functioning. Assessment should allow for the measurement of both subject matter learning (e.g. quantity and quality of contribution) and teamwork learning (e.g. interpersonal communication and cooperation). Such attributes are more amenable to peer rating rather than peer ranking methods of evaluation (Lejk & Wyvill, 2001).

With the learning of teamwork skills as an objective of group based assignments, it stands to reason that peer evaluations could aid in that process by serving not only the purpose of a summative evaluation for grading, but also as formative evaluation and feedback to the student. If an evaluation is to serve as feedback, then students need to receive the results of their peer evaluation and a mechanism needs to be provided for discussion. Interestingly, however, peer evaluations in the literature seem to be used largely with the data collected anonymously for incorporation into marks (Asgari & Dall'Alba, 2011; Bacon, Stewart & Silver, 1998; Lejk & Wyvill, 2001; Saito & Fujita, 2009) at the conclusion of a project, rather than for provision to the student to facilitate his or her learning. Some students indicate that such use can leave them feeling blindsided and cheated of the opportunity to improve their performance had they received direct feedback from team members (Feitchner & Davis, 1984). In those cases where peer evaluation is seen as a potential source of feedback, the evaluations are still often collected anonymously (e.g. Brutus &

Donia, 2010), thereby inhibiting learning about how to have an honest and open discussion amongst team members about their process.

The emphasis on anonymity of peer evaluation seems to stem in part from a concern that friendship bias will lead to a lack of honesty amongst peers (Barclay & Harland, 1995). Indeed, some students report that there is peer pressure to provide positive feedback (Burdett, 2007). Research also suggests, however, that the discussion of peer evaluations early in group work gives greater clarity as to what constitutes high quality work (Topping, 1998), allows for the identification of potential issues (Burdett, 2007), and has a positive impact on perceptions of cohesion and satisfaction in group work (Druskat & Wolff, 1999; Feitchner & Davis, 1984). If peer evaluations are non-confidential, they can also facilitate the learning of how to provide constructive criticism and how to work through the issues that can be so destructive to group functioning.

The dilemma, therefore, is: in order to solve one problem, which is the lack of a vehicle for honest and open feedback to students regarding their functioning in a group, do we as instructors exacerbate another problem, which is the social loafing and inequitable mark distribution? By making peer evaluations non-confidential, we may facilitate student learning about how to work in a team, but we may also inflate individual marks for students who did not contribute their share. The question as to if the confidentiality of peer evaluation will alter its impact on course marks has not been investigated in the literature to date. In fact, some research gives no indication of if the evaluations themselves were anonymous (e.g. Cheng & Warren, 2000 etc.), suggesting that perhaps this issue has not been considered. In this article, we aim to investigate the impact of confidentiality on peer ratings of individuals' team work, in particular those poor attending, social loafing and domineering behaviors associated with group derailment, and the subsequent grading of individuals on group projects. If the lack of confidentiality recommended to facilitate the learning of team-based skills does lead to mark inflation, then faculty would need to incorporate strategies to adjust for this issue in order to ensure fair mark distribution within the group.

Method

Participants and Procedures

Eighty-six students enrolled in one of four sections of a three-credit *Introduction to Research Methods* course in a master's-level education program in the US participated in this study. As part of the course, each student was assigned to a group of between 3 and 4 students. The groups worked together for the entire semester to complete two projects: a literature review and a survey research study on a topic of interest. Marks were assigned on the basis of the final two products and the peer evaluation. The sections were randomly assigned to the comparison or experimental group, with two sections in the comparison group and two in the experimental group.

In order to control for threats to internal validity, the procedure in both groups was equivalent with the only difference being that of the confidentiality of the peer evaluation. Both groups were told at the beginning of the semester that they would have to complete a peer evaluation of their group mates at the conclusion of the semester. The evaluation form was provided for their review and for questions and they were told that each person's peer evaluations would be incorporated into his or her final mark. Both groups completed the peer evaluation form at the end of the semester. Students in the comparison group were

told that their ratings and comments would be shared only with the professor and not with the evaluatees. Students in the experimental group were told that their evaluations would be shared with the evaluatees and that the purpose of the evaluation was twofold: 1) to ensure that the individual's project mark reflected their contribution and 2) to provide their colleagues with some feedback on what they did well and what they could improve in their group work. All peer evaluation ratings and comments were compiled by the professor and were provided to evaluatees in the experimental group via email at the conclusion of the semester.

There were 47 students in the experimental group with 96% female, 4% male. In the comparison group there were 39 students, of which 79% were female, 21% male.

Measures

Peer Evaluation

In the peer evaluation, students used a peer-rating evaluation form to rate their colleagues on five characteristics for a total of 50 points. Each characteristic was measured via a 10-point rating scale with behavioral anchors at either extreme and in the middle (Appendix). The scale was designed to tap attributes of three team-member types frequently cited in the literature as sources of student group derailment: the poor attender, the domineering teammate, and the social loafer. The attributes, when lacking, associated with the team member type were as follows: items 1 and 2, attendance and preparation, were characteristics of the poor attender, item 3, contribution, was associated with the social loafer and items 4 and 5, respect for others' ideas and flexibility, was associated with the domineering teammate. The individual student's total peer evaluation score was the average of his or her evaluations from all teammates.

Students were also asked to provide comments and an explanation for each low rating on the peer evaluation form. When evaluations were confidential, this allowed the professor to assess the thinking of the evaluator. When evaluations were non-confidential, this provided critical feedback to the evaluatee.

Group Project Subject Marks

Groups received a mark from 0-100 for each of two group projects: a literature review and a survey research project. The overall group mark was the average of these two group projects. For the purposes of this research, the peer evaluation results were incorporated into the individual's overall group mark via two methods: an additive method (Maranto & Gresham, 1998) and a weighted method (Lejk & Wyvill, 2001; Zhang & Ohland, 2009). In the additive method, the individual's average peer evaluation score was added to the group mark and computed as a percentage of the total possible points for the individual's final score. Conceptually, this can be represented as:

$$Final\ Mark_{Additive} = \frac{Individual's\ Group\ Mark + Peer\ Evaluation}{Maximum\ Possible\ Group\ Mark + Peer\ Evaluation}$$

Using an example, a student who scored a group mark of 90 and a peer evaluation score of 45 would have a combined score of 90+45=135 divided by the total possible points (150). This would result in a final additive mark of 0.9 or 90%.

If, as suggested in the literature, a lack of anonymity may lead to friendship bias in student ratings (Barclay & Harland, 1995; Burdett, 2007), then the additive method will lead to higher marks for students in the experimental group as compared to the comparison group. For this reason, we also computed marks using a weighted method (Zhang & Ohland, 2009)

which does not use a raw peer evaluation score, but rather creates a weighted peer evaluation score relative to the peer evaluation scores of his or her team members. This weighted score is derived by dividing the individual's peer evaluation score by the average peer evaluation score of the entire group, as follows:

$$\text{Weighted Peer Evaluation Score} = \frac{\text{Individual's Peer Evaluation Score}}{\text{Group Average Peer Evaluation Score}}$$

Conceptually, this weighted individual evaluation score represents how the rating of the individual compares to the average rating of the group. Students whose peer evaluation score is higher than the average peer evaluation score in the group will have a weighted score that is greater than 1. Those whose peer evaluation score is lower than the average peer evaluation score will have a weighted evaluation score less than one, and if their peer evaluation score is the same as those of the others in the group, then the weighted score will be 1. The student's final mark, then, is calculated by multiplying the weighted peer evaluation score by the group mark, thusly:

$$\text{Final Mark}_{\text{Weighted}} = (\text{Group Mark}) \text{Weighted Peer Evaluation Score}$$

Using the above example, if the student who had a peer evaluation score of 45 was in a group where the two other team members had a peer evaluation score of 48 each, then the average peer evaluation score in that group would be 47. The student with a 45 would have a weighted peer evaluation score of $45/47$ or .96. Conceptually this represents the contribution of this individual relative to his or her other group members. This individual's final mark would then be a proportion of the group's mark based on his contribution as: $90(.96) = 86$.

Results

Prior to testing our main research question, we compared the two groups (i.e., confidential versus non-confidential) for the potentially confounding factors of gender and general ability. As research suggests that peer evaluators evaluate males and females differently (Schonrock-Adema, Hijne-Penniga, van Duijn, Geertsma & Cohen-Schotanus, 2007), we wanted to be sure that the two groups were roughly equivalent in terms of their numbers of males and females. The confidential peer evaluation group had a significantly greater percentage of males (20.5%) than the non-confidential peer evaluation group (4.3%) ($X^2(1) = 5.48, p < .05$). In addition, as each individual's peer evaluation scores were combined with his or her overall group marks (i.e., the average of the two group projects), we wanted to be sure that the two groups did not differ on their overall group marks prior to the introduction of the peer evaluation scores. There was no statistically significant difference between the groups on overall group marks ($F(1, 83) = .007, p = .94$), suggesting that they were roughly equivalent in general ability and rendering statistical control unnecessary. Because the two groups differed in terms of the number of males, however, gender was statistically controlled for in subsequent analyses.

To test the assumption that non-confidential peer evaluation would result in mark inflation, we compared the differences between the students' marks without the peer evaluation (the overall group mark) to the students' marks with the peer evaluation added using the additive method (final mark_{additive}) for the two groups (i.e., confidential versus non-confidential), controlling for gender (Table 1). Results were as expected for the non-

confidential peer evaluation group; the students' overall group mark without the peer evaluation was significantly lower than the students' final mark_{additive} with the peer evaluation. This supports the assertion that the elimination of confidentiality results in higher peer ratings and, consequently, higher overall marks. Interestingly, among the confidential peer evaluation group, the students' overall group mark was significantly higher than the final mark_{additive}, suggesting that the assurance of confidentiality results in lower peer ratings and, consequently, lower overall marks.

Table 1.

Two-Way Mixed Analysis of Covariance (ANCOVA) Comparing Confidentiality Groups with Peer Evaluations Added to Overall Group Mark Using the Additive Mark Method¹

Group	Overall Group Mark Mean	Final Mark _{Additive} Mean
Confidential	92.06 ^a	90.91 ^a
Non-Confidential	92.12 ^b	93.02 ^b

¹ Significant interaction effect, Wilks' Lambda=.84; $F(1,83)=15.63$, $p<.001$ for the 2 (Group: Confidential versus Non-Confidential) X 2 (Mark Method: Overall Group Mark versus Final Mark_{Additive}) mixed ANCOVA

^a Significantly different, $p < .05$

^b Significantly different, $p < .05$

Finally, we examined if the weighted method of computing marks for the two peer evaluation groups would correct for the differences encountered above (Table 2). If the weighted method does correct for the confidentiality effect, then we should see no differences between the students' marks without the peer evaluation and the student's marks with the peer evaluation added in using the weighted method for either group. The results indicated that the interaction effect was not statistically significant. In other words, the overall group mark without the peer evaluation was not significantly different from the final mark with the peer evaluation added using the weighted method (final mark_{weighted}) for the confidential peer evaluation group nor for the non-confidential peer evaluation group.

Table 2

Two-Way Mixed Analysis of Covariance (ANCOVA) Comparing Confidentiality Groups with Peer Evaluations Added to Overall Group Mark Using the Weighted Mark Method²

Group	Overall Group Mark Mean	Final Mark _{Weighted} Mean
Confidential	92.06	92.33
Non-Confidential	92.12	91.89

² Non-significant interaction effect, Wilks' Lambda=.99; $F(1,83)=.04$, $p=.83$ for the 2 (Group: Confidential versus Non-Confidential) X 2 (Mark Method: Overall Group Mark versus Final Mark_{Weighted}) mixed ANCOVA

Discussion

Peer evaluation has been proposed as a vehicle to facilitate group discussion on group functioning issues (Brutus & Donia, 2010) and to differentiate the marking of individuals working in teams (Cheng & Warren, 2000; Zhang & Ohland, 2009). It is possible, however, that these two purposes may be at odds with one another. In order for peer evaluations to

be a tool for students to address group process, we argue that they should be conducted in an open and transparent manner, with the results presented to the evaluatee for feedback. Peer pressure leading to elevations in peer evaluation scores and subsequent marks when the process is not anonymous (Barclay & Harland, 1995; Burdett, 2007), however, might render useless the impact of peer evaluation on differentiating student marks within a group.

Our results show that concerns about mark inflation when using non-confidential evaluations are warranted, as when non-confidential peer evaluations were combined additively with group marks, the average mark was significantly higher than without the peer evaluations. Interestingly, confidentiality does not appear to resolve this problem, as when confidential peer evaluations were added to group marks, the average mark was significantly lower than group marks. This suggests that, regardless of confidentiality type, merely adding peer evaluation scores into an individual's group marks will cause a significant shift for the entire class in one or the other direction. Such a shift would not correct for the inequitable mark distribution so often cited by students as a source of their hostility toward group work. Faculty who decide to use peer evaluations as a mechanism to assess individual performance in a group, therefore, will have to use caution in incorporating evaluation scores into the student marks.

Our findings suggest that the weighted method proposed by Zhang and Ohland (2009) to combine peer evaluation scores with group marks corrects for both the inflation and deflation effects of confidentiality type. Although individual student marks change due to their weighted peer evaluation scores, they do not all change in the same direction and the overall mean does not differ significantly from the mean overall group mark without the inclusion of the peer evaluation scores. We propose that the weighted method may more effectively correct for the mark distribution inequity by ensuring that students' contributions are directly compared to those of their own group members. Those with lower ratings, and, possibly, lower contributions, receive a lower proportion of the group mark than their more highly rated group members. If an issue of concern is inequitable mark distribution, then the weighted method of combining peer evaluations with the group mark looks promising.

The overall implication is that a) confidentiality impacts student marks and b) regardless of confidentiality type, faculty must correct for this impact via a weighted marking method. For those with fewer students, the weighted method would simply require the addition of one computation per group (the average rating per group) and per individual (the weighted peer evaluation score) and the inclusion of this data point in the computation of the overall mark. For faculty working with larger numbers of students and groups, such a method may be time prohibitive. In such cases, faculty may want to consider eliminating peer evaluations from their marking procedures so as to avoid the deflation of marks associated with confidential evaluations and the inflation association with non-confidential evaluations.

When peer evaluation is an option, we recommend an evaluation form designed to assess for the social loafing, domineering behaviors, and poor attendance so often cited by students as sources of difficulty in their group work. In addition, the use of non-confidential peer evaluations to facilitate the discussion and problem-solving so essential to the development of team skills is recommended. It should be noted that the use of non-confidential evaluations at the end of the semester as described herein enabled the authors to compare the impact of confidentiality and non-confidentiality on marks but would be insufficient to facilitate the rich learning about how to work in teams that we are suggesting is critical in team-based assignments. For such feedback to be of use, students would need

to receive it midway through the group's work together and a vehicle would need to be provided for processing and discussion.

Some research suggests that team-based skills do not develop in student groups without faculty intervention (Brutus & Donia, 2010; Hertz-Lazarowitz, 1989), but faculty support for student groups typically occurs only in the form of specific instructions regarding the academic task (Hansen, 2006) with little, if any, instruction provided on how to develop and use teamwork skills (Colbeck, Campbell, & Bjorklund, 2000; Myers et al., 2009). It has been argued that faculty often are unfamiliar themselves with team process strategies and feel ill equipped to coach students in managing the emotion associated with incidences of social loafing, domineering behaviors, or poor attendance (Burdett, 2007). The group emotional intelligence literature can provide guidance in this area. Strategies for developing a group's emotional intelligence have been proposed by a number of researchers and could serve as models for intervention (e.g. Amundson, 2005; Clinebell & Stecher, 2003; Druskat & Wolff, 2001; Kremenitzer, Mojsa, & Brackett, 2008).

Finally, there were several limitations to this study. First, the sample was comprised of graduate students, largely women, in a private school of education. This may limit generalizability to younger, more diverse student populations. In addition, although this study explored the impact on marks of non-confidential peer evaluations, additional research must be conducted to determine if such evaluations actually foster discussion of group issues and, subsequently, reduce domineering, social loafing, and poor attendance behaviors.

References

- Amundson, S.J. (2005). The impact of relational norms and the effectiveness of health and human service teams. *The Health Care Manager, 24*(3), 216-224.
- Ashraf, M. (2004). A critical look at the use of group projects as a pedagogical tool. *Journal of Education for Business, 79*(4), 213-216.
- Asgari, S. & Dall'Alba, G. (2011). Improving group functioning in solving realistic problems. *International Journal for the Scholarship of Teaching and Learning, 5*(1), 1-14.
- Bacon, D.R., Stewart, K.A. & Silver, W.S. (1999). Lessons from the best and worst student team experiences: How a teacher can make the difference. *Journal of Management Education, 23*(5), 467-488.
- Barclay, J.H. & Harland, L. (1995). Peer performance appraisals: The impact of rater competence, rater location, and rating correctability on fairness perceptions. *Group and Organization Management, 20*(1), 39-60.
- Boud, D., Cohen, R. & Sampson, J. (1999). Peer learning and assessment. *Assessment & Evaluation in Higher Education, 24*(4), 413-426.
- Brutus, S. & Donia, M.B.L. (2010). Improving the effectiveness of students in groups with a centralized peer evaluation system. *Academy of Management Learning & Education, 9*(4), 652-662.

- Burdett, J. (2007). Degrees of separation – balancing intervention and independence in group work assignments. *The Australian Educational Researcher*, 34,(1), 55-71.
- Bushell, G. (2006). Moderation of peer assessment in group projects. *Assessment & Evaluation in Higher Education*, 31(1), 91-108.
- Campion, M.A., Papper, E.M., & Medsker, G.J. (1996). Relations between work team characteristics and effectiveness: A replication and extension. *Personnel Psychology*, 49, 429-452.
- Carlsmith, K.M. & Cooper, J. (2002). A persuasive example of collaborative learning. *Teaching of Psychology*, 29(2), 132-135.
- Carson, K.M. & Glaser, R.E. (2010). Chemistry Is in the News: Assessing intra-group peer review. *Assessment & Evaluation in Higher Education*, 35(4), 381-402.
- Cheng, W. & Warren, M. (2000). Making a difference: Using peers to assess individual students' contributions to a group project. *Teaching in Higher Education*, 5(2), 243-255.
- Clinebell, S. & Stecher, M. (2003). Teaching teams to be teams: An exercise using the Myers-Briggs Type Indicator and the five-factor personality traits. *Journal of Management Education*, 27(3), 362-383.
- Colbeck, C.L., Campbell, S.E. & Bjorklund, S.A. (2000). Grouping in the dark: What college students learn from group projects. *The Journal of Higher Education*, 71(1), 60-83.
- Conway, J.M. & Huffcutt, A.I. (1997). Psychometric properties of multisource performance ratings: A meta-analysis of subordinate, supervisor, peer, and self-ratings. *Human Performance*, 10(4), 331-360.
- Druskat, V.U. and Wolff, S.B. (1999). Effects and timing of developmental peer appraisals in self-managing work groups. *Journal of Applied Psychology*, 84(1), 58-74.
- Druskat, V.U. and Wolff, S.B. (2001). Building the emotional intelligence of groups. *Harvard Business Review*, 80-90.
- Falchikov, N. & Goldfinch, J. (2000). Student peer assessment in higher education: A meta analysis comparing peer and teacher marks. *Review of Educational Research*, 70(3), 287-322.
- Feichtner, S.B. & Davis, E.A. (1984). Why some groups fail: A survey of students' experiences with learning groups. *The Organizational Behavior Teaching Review*, 9(4), 58-73
- Hansen, R.S. (2006). Benefits and problems with student teams: Suggestions for improving team projects. *Journal of Education for Business*, 82(1), 11-19
- Harkins, S.G. and Jackson, J.M. (1985). The role of evaluation in eliminating social loafing. *Personality and Social Psychology Bulletin*, 11(4), 457-465.
- Hertz-Lazarowitz, R. (1989). Cooperation and helping in the classroom: A contextual approach. *International Journal of Research in Education*, 13, 113-119.

- Hoffman, J.R. & Rogelberg, S.G. (2001). All together now? College students' preferences in project group grading procedures. *Group Dynamics: Theory, Research and Practice*, 5(1), 33-40.
- Jalagas, D.S. & Sutton, R.I. (1984). Feuds in student groups: Coping with whiners, martyrs, saboteurs, bullies, and deadbeats. *The Organizational Behavior Teaching Review*, 9(4), 94-102.
- Johnson, D.W. & Johnson, R.T. (1992). Positive interdependence: Key to effective cooperation. In R. Hertz-Lazarowitz & n. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 174-199). New York: Cambridge University Press.
- Johnson, D. W., & Johnson, F. P. (2008). *Joining together: Group theory and group skills* (10th ed.). Pearson.
- Johnson, D.W., Johnson, R.T. & Smith, K.A. (1998). Cooperative learning returns to college. *Change*, 30(4), 26-35.
- Kagan, S. (1995). Group grades miss the mark. *Educational Leadership*, 52(8), 68-71.
- 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.
- Kremenitzer, J.P., Mojsa, J.K. & Brackett, M.A. (2008). Creating an emotionally intelligent classroom culture. In R.J. Emmerling, V.K. Shanwal, & M.K. Mandal (Eds.) *Emotional Intelilgence: Theoretical and Cultural Perspectives* (pp. 191-207). Hauppauge NY: Nova Science Publishers.
- Latane, B., Williams, K. & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, 37(6), 822-832.
- Lejk, M. & Wyvill, M. (2001). Peer assessment of contributions to a group project: A comparison of holistic and category-based approaches. *Assessment & Evaluation in Higher Education*, 26(1), 61-72.
- Li, L.K.Y. (2001). Some refinements on peer assessment of group projects. *Assessment & Evaluation in Higher Education*, 26(1), 5-18.
- Longmore, M.A., Dunn, D., & Jarboe, G.R. (1996). Learning by doing: Group projects in research methods classes. *Teaching Sociology*, 24(1), 84-91.
- Maranto, R. & Gresham, A. (1998). Using "World Series Shares" to fight free riding in group projects. *PS: Political Science and Politics*, 31(4), 789-791.
- Meyers, S.A. (1997). Increasing student participation and productivity in small-group activities for psychology classes. *Teaching of Psychology*, 24(2), 105-115.
- Monk-Turner, E., & Payne, B. (2005). Addressing issues in group work in the classroom.

Journal of Criminal Justice Education, 16, 166-179.

Myers, S.A., Smith, N.A., Eidsness, M.A., Bogdan, L.M., Zackery, B.A., Thompson, M.R., Schoo, M.E., & Johnson, A.N. (2009). Dealing with slackers in college classroom work groups. *College Student Journal*, 43(2), 592-598.

O'Brien, K.M. (1995). Enhancing research training for counseling students: Interuniversity collaborative research teams. *Counselor Education and Supervision*, 34(3), 187-198.

Peterson, S.E. & Miller, J.A. (2004). Comparing the quality of students' experiences during cooperative learning and large-group instruction. *The Journal of Educational Research*, 97(3), 123-133.

Pope, N. (2001). An examination of the use of peer rating for formative assessment in the context of the theory of consumption values. *Assessment & Evaluation in Higher Education*, 26(3), 235-246.

Revere, L., Elden, M. & Bartsch, R. (2008). Designing group examinations to decrease social loafing and increase learning. *International Journal for the Scholarship of Teaching and Learning*, 2(1), 1-16.

Saito, H. & Fujita, T. (2009). Peer-assessing peers' contributions to EFL group presentations. *Regional Language Centre Journal*, 40(2), 149-171.

Schonrock-Adema, J., Heijne-Penniga, M., van Duijn, M.A.J., Geertsma, J. & Cohen Schotanus, J. (2007). Assessment of professional behavior in undergraduate medical education: Peer assessment enhances performance. *Medical Education*, 41, 836-842.

Slavin, R. (1995). *Cooperative learning: Theory, research and practice* (2nd edition). Baltimore, MD: Johns Hopkins University.

Su, A.Y. (2007). The impact of individual ability, favorable team member scores, and student perception of course importance on student preference for team-based learning and grading method. *Adolescence*, 42(168), 805-826.

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249-276.

Zhang, B. & Ohland, M.W. (2009). How to assign individualized scores on a group project: An empirical evaluation. *Applied Measurement in Education*, 22, 290-308.

Appendix: Team Member Assessment

Name of team member: _____

Rate the team member named above on each of the following five scales. A maximum score for each scale is 10 and a minimum score is 1.

Scale	Score
Attendance	
Missed several meetings/classes, was frequently late or left early	Attended all meetings/classes from start to finish
Missed or was late to a few meetings/classes	
1 2 3 4 5 6 7 8 9 10	
Preparation	
Had not read materials and prepared before meetings/class	Had read all materials and prepared thoroughly before meetings/class
Was sometimes prepared for meetings/class	
1 2 3 4 5 6 7 8 9 10	
Contribution (Quality and Quantity)	
Contributed little and what was produced was low quality	Contributed substantially and produced high quality work
1 2 3 4 5 6 7 8 9 10	
Respect for Others' Ideas	
Did not listen to and/or scoffed at others' ideas	Actively sought others' input, listened well, spoke respectfully to others
1 2 3 4 5 6 7 8 9 10	
Flexibility	
Was uncompromising and inflexible	Orchestrated effective compromises when disagreements occurred
Was willing to adapt to new ideas but needed help figuring out how	
1 2 3 4 5 6 7 8 9 10	
TOTAL SCORE OUT OF 50 POSSIBLE POINTS	

Provide comments here and on the back of the page: