A TECHNIQUE FOR PEER AND GROUP MEMBER EVALUATIONS OF TEAMWORK IN THE UNDERGRADUATE CLASSROOM

Ranida B. Harris, Indiana University Southeast, School of Business, rbharris@ius.edu
Kenneth J. Harris, Indiana University Southeast, harriskj@ius.edu
Kara L. Vaught, Indiana University Southeast, kavaught@ius.edu

ABSTRACT

As the information systems (IS) world in which business occurs continues to change, teams are being utilized with even greater frequency. As a result, and to better prepare students for the working world, IS educators often utilize teamwork in their classes. Although benefits from working with others are widely known, there are also negative aspects of teamwork. One such problem relates to how individual efforts can be fairly and accurately evaluated in a team so that all group members do not receive the same grade unless they have actually earned it. This paper discusses a number of different issues related to current teamwork evaluations and provides an assessment tool that calculates an individual's grade by using the overall group grade and group member evaluations relative to the evaluations received by their other group members. Observed outcomes and feedback received are discussed.

Keywords: Evaluation, Teamwork, Instructional Materials, IS Education

INTRODUCTION

For years, information systems (IS) educators have been increasingly utilizing teams for class assignments and projects. The reasons for employing teamwork are varied, with one of the greatest reasons being the desired learning outcomes. In the workplace, IS work is often structured around teams, thus making it all the more important that students graduate having experience with and knowledge of teamwork. Additionally, numerous research efforts have shown that employers value communication skills, team problem solving, being able to work well with others, and other interpersonal skills that can be improved through teamwork [3, 10, 11].

Another reason for the use of teams in IS classes relates to necessity. As budgets have tightened and class sizes have become larger, it becomes less realistic for an instructor to assign and give thorough, meaningful feedback on individualized projects. Further, as IS educators strive to have more realistic and challenging projects that better portray real-world activities, they often find that the projects require more time than a single student would be able to provide [2]. Thus, educators utilize teams as they are able to collectively complete the large-scale projects that individual students find overwhelming [20].

For these and other reasons, IS educators often utilize teams to complete class projects. However, as with the structure of a considerable number of assignments, there are pros and cons to having students work in teams. One disadvantage in particular is that there are often problems with evaluating and rewarding individual efforts in the context of teamwork. Although numerous peer assessment tools are currently available, many are flawed. This paper hopes to aid IS educators by providing details on a peer evaluation assessment tool that incorporates into the final grade the components of the overall team grade, the average peer member evaluation, and the relative peer member evaluation compared to other group members. To help demonstrate the need for this teamwork evaluation technique, the following sections discuss the benefits of teamwork in classes as well as problems with groupwork and the current assessment tools.

BENEFITS OF TEAMWORK IN CLASSES

As previously mentioned, work is often structured around teams in the real world. One of the primary motivations for these job designs rests in the fact that teamwork has been shown to result in a number of positive outcomes including a greater pool of knowledge, better acceptance of the solution, and a higher quality product [6]. Because teams are a part of the IS work environment, it becomes even more crucial for IS classes to involve teamwork.

Through working in teams, students are able to learn teamwork skills such as group communication, brainstorming, project management, synchronous and asynchronous collaboration, task coordination, and more [9]. Additionally, students often learn important management skills including how to handle disruptions and team conflict. Further, there are
often situations in which one or more team members fail to perform their portion of the work assigned to a team. Thus, team members get the chance to work on skills including motivation, engaging in critical conversations, and ultimately, potentially completing the work for someone else.

**DISADVANTAGES OF TEAMWORK IN CLASSES**

Unfortunately, with teamwork, there are also various disadvantages for both students and IS educators. For students, teamwork is often associated with a number of disadvantages, including increased conflict, greater difficulty in coordinating and completing the work, dominance by certain group members, and social pressures to go along with the majority. Other problems include social loafing [14] and free riding [1], which refer to the fact that team members often expend less effort when the team size is larger and often try to benefit from the contributions that other members make while providing minimal contributions. When there are differences in team member contributions but students receive the same grades, those who provided increased efforts are frequently left dissatisfied with the overall experience.

The problems of social loafing and free riding for team members lead to one of the primary disadvantages for instructors, as it is difficult to accurately assess and grade the inputs of individuals working in teams. Additionally, if student efforts are different but all team members receive the same grade, the top performers are likely to perceive this process as unfair and be less motivated to work hard on future team projects [15]. On the flipside, if social loafers and free riders receive the same grade as those who contributed considerably more on the project, then they are likely to feel as though they “won” and believe that working hard on group work is not really necessary.

Thus, there is a need for IS educators to have assessment tools that not only accurately access individual team member inputs, but also reward (grade) them accordingly. In the sections that follow, we discuss a number of evaluation tools, the problems with those techniques, and a new peer member assessment and grading tool to be utilized with group assignments.

**CURRENT TEAM ASSESSMENT TOOLS AND THEIR WEAKNESSES**

Due to the aforementioned potential issues related to the assessment of teamwork, there is a need for instructors to evaluate and assign final grades for each team member of a group. A practice used in a large number of cases involves the instructor evaluating a project and assigning the same grade to all individuals in the group. Although this assessment method is easy, it has a number of drawbacks. For one, there is the problem related to social loafing and free riders. Additionally, when students know that they are only evaluated as a group, not individually, and they know they have a hard-working group member or two, they are even less likely to put forth their fair share of the work (i.e., “pull their own weight”). Secondly, this evaluation technique does not differentiate between team members. Thus, if some students contribute more and some contribute less to the final product, they all receive the same grade, which is not an accurate representation of their efforts. Further, studies have shown that students feel that the method of assigning an overall project grade with no team member differences is not a fair evaluation of each person’s efforts[5]. Finally, if all group members receive the same grade on a team project, there is likely to be grade clustering and a “non-normal” distribution of overall class grades.

Because of the problems mentioned above, a number of educators and research efforts have been focused on how to best evaluate individual performance when students work in teams [7, 8, 12, 16, 17, 18, 22]. This is necessary because not only do students work with different levels of intensity, diligence, and quality, but they also bring different skills (e.g., ability, experience, knowledge) to the project that may result in differential inputs which need to be evaluated accordingly. However, there is a problem due to asymmetric information [22]. In layman’s terms, this means that although the team members know the contributions of themselves and others in the team, the instructor does not have access to this information.

As a result, numerous instructors and research efforts have focused on some form of peer member evaluations [7, 17, 22]. However, the traditional approaches often involve students rating their group members on an overall assessment (e.g., overall performance), and the average of their group member ratings is used as a certain component of the overall grade. Unfortunately, this approach, which is the
basis for a large number of similar peer evaluation techniques, is subject to a number of flaws. These problems are shown in Table 1 and further described below.

Table 1
Problems with the traditional peer member evaluation technique

<table>
<thead>
<tr>
<th>Problem</th>
<th>Description of the Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade Inflation</td>
<td>Students give group members higher grades than the overall professor grade, resulting in overall grades increasing</td>
</tr>
<tr>
<td>Lack of Differentiation</td>
<td>Students don’t want to be critical, or they give everyone the same grade (often 100’s), which minimizes the usefulness of this technique</td>
</tr>
<tr>
<td>Only an Overall Performance Criteria</td>
<td>Although performance within a team has many components, many evaluation tools only assess overall performance</td>
</tr>
<tr>
<td>Students “Playing the Game”</td>
<td>Students often make deals to give everyone high grades, even if the grades are not deserved</td>
</tr>
<tr>
<td>Non-objective/No Training</td>
<td>In many cases students don’t have to provide objective comments, and they lack the training to provide high-quality comments</td>
</tr>
</tbody>
</table>

One of the primary problems with having students complete peer evaluations is that they often result in grade inflation. For example, we can examine a situation in which the instructor assigns the overall project grade by allocating 75% to the instructor group grade and 25% to the average of peer evaluations. If the group receives a grade of 85 on the project but the average peer evaluations are 95, then the average student grade will be an 87.5 \((0.75*85) + (0.25*95) = 87.5\), even though the group only received an 85 on the assignment. Some reasons that peer evaluations are often higher than instructor grades is that students are generally not as critical and they tend to find it difficult and risky to negatively rate a peer member or friend [13]. A related problem is that students often fail to differentiate when completing their peer evaluations. The reasons are similar, if not identical; students often think it is easier, they want to be nice, or they want to get finished quicker.

A third problem with a number of the current evaluation methods is that they only assess overall performance. Although this may be appropriate in certain cases, it fails to accurately reflect reality, as performance is multi-dimensional. A fourth problem relates to the fact that some students may make deals to give each other higher scores. In these cases, if students know that they can help each others’ grades and/or it makes completing the project easier (e.g., if experiencing conflict, everyone agrees to give each other high grades and that lessens tension), they may do so, which again decreases the usefulness of the peer evaluations.

A final problem with most peer member evaluation tools is that students often doubt the objectivity of these forms, as comments are generally not required and students have no training in assessment [4, 19]. Students feel that the evaluation forms do not allow them to explain themselves, and even if they do, they are full of subjective comments. Additionally, for most forms that do include comments, students aren’t trained in how to use them, what kinds of comments to include, etc.

THE TEAMWORK ASSESSMENT UTILIZED

As a result of the aforementioned problems with existing peer evaluation forms, we set out to develop an assessment tool that would maintain the benefits derived from peer evaluations, but eliminate or minimize as many of the problems as possible. The evaluation tool, which is provided in Appendix 1, will be explained. Additionally, the formula involved in calculating the student grades is a vital part of the evaluation tool, and will also be explained to help clarify the team member assessment tool.

First, as can be seen, the evaluation form is designed to assess the multi-dimensional nature of performance, and does so by examining the different components of “Quality of Work,” “Quantity of Work,” and “Being a Good Team Member.” Second, students are informed in class that written (or typed) comments are important in defending the grades provided. Further, they are provided instructions on how to complete the form, and are advised that they should attempt to use objective comments to back up their statements as best possible. The idea behind this is that objective comments are less subjective and much closer to facts. Third, they are informed that they have to provide self-ratings, which are not used in the formula, but are used to determine if someone can accurately assess themselves and if they are overly harsh or easy when assessing their peers. Finally, the students are told to be as accurate as possible and remember that if they are assigning a
grade for someone who didn’t really earn it, they are only reinforcing a problem. We encourage students to think of the situation from an ethical standpoint.

Once all group members have completed the evaluation form, it is time to utilize the formula, which will be described below. In the syllabus, students are informed that for group projects, 50% of the grade will come from the instructor’s evaluation and 50% will come from relative peer member evaluations, with the overall grade based on the instructor’s evaluation. The idea here is that by using the instructor’s initial grade for the project as the basis for how peer member evaluations impact the grade, the overall grade within a group will not be increased (i.e., eliminating problems of grade inflation). In layman’s terms, this would mean that if a group of four members received a group grade of 80, and person A had peer evaluation averages for the three criteria of 100, person B had an average of 80, person C had an average of 80, and person D had an average of 60, then person A’s grade for the peer evaluation component would go up by X points, persons B and C would have their portion of the peer evaluation component remain a 40 (80 * 50%), and person D’s grade would go down by X points. This formula will be described in greater detail below, but as can be seen, the overall average in this case stays at the overall group grade.

For an illustration of the point breakdown using the evaluation tool and formula described, two examples are provided below. In example 1, the group grade from the instructor was a 90 and the average from all three dimensions of performance rated (quality of work, quantity of work, and being a good team member) from all group members was a 100 for group member A, a 90 for group member B, an 80 for group member C, and a 70 for group member D. As can be seen in the table below, 50% of the grade comes directly from the instructor grade (50% * 90), and the other 50% of the grade comes from the relative peer member evaluations. The 50% of the grade that comes from the peer member evaluations is calculated by taking the totals received for each person (300 for group member A, 270 for group member B, 240 for group member C, and 210 for group member D) and adding all of those totals together (300 + 270 + 240 + 210 = 1020). Then, each group member’s total peer evaluation number is divided by the overall number, which would result in 25% if everyone received equal peer member evaluations. The resulting number is multiplied by 2 to get to 50%, and then multiplied by the instructor-given overall grade of 90 to receive the other 50% of the grade. As can be seen, after the peer member evaluations are calculated, group member A, who received the highest peer evaluations, received a grade of 97.94, group member B’s overall project grade was 92.64, group member C’s overall grade was 87.36, and group member D’s was 82.06. Thus, from this example it can be seen that there is differentiation in grades, individuals are rewarded for extra effort and punished for below-average effort, and there is no grade inflation.

### Example 1: Group Grade with Peer Evaluations Factored In

- Group has 4 people: A, B, C, and D and receive a group grade of a 90
- Each group member rates themselves and their group members on the 3 dimensions. Then the instructor calculates the average for the 3 components for each group member.
- The average ratings for each group member are: A=100, B=90, C=80, and D=70
- 50% of the grade comes from the instructor grade, and 50% is determined from the relative peer member evaluations (as shown below)

<table>
<thead>
<tr>
<th>Member</th>
<th>Rater</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>Total Eval.</th>
<th>Average Eval.</th>
<th>Final Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td>X</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>300</td>
<td>100</td>
<td>(90 * 50%) + (300/1020 * 2 * 90) = 97.94</td>
</tr>
<tr>
<td>B</td>
<td>100</td>
<td>X</td>
<td>80</td>
<td>90</td>
<td>270</td>
<td>90</td>
<td>90</td>
<td>(90 * 50%) + (270/1020 * 2 * 90) = 92.64</td>
</tr>
<tr>
<td>C</td>
<td>100</td>
<td>65</td>
<td>X</td>
<td>75</td>
<td>240</td>
<td>80</td>
<td>80</td>
<td>(90 * 50%) + (240/1020 * 2 * 90) = 87.36</td>
</tr>
<tr>
<td>D</td>
<td>100</td>
<td>55</td>
<td>55</td>
<td>X</td>
<td>210</td>
<td>70</td>
<td>70</td>
<td>(90 * 50%) + (210/1020 * 2 * 90) = 82.06</td>
</tr>
</tbody>
</table>

Total: 1020 Average ~ 90

Another illustration can be seen in example 2. In this case, the group received an average grade of 85 and then all group members gave each other peer member evaluations of 100. If this example took place using an evaluation method in which the instructor-provided grade was 50% and the peer evaluations accounted for 50%, each group member would end up with a grade of 92.5 ((50% * 85) + (50% * 100) = 92.5). This is a simple illustration of grade inflation and how grades could very easily increase if all group members gave each other high evaluations. On the other hand, using the peer evaluation tool and formula described in this paper and shown in example 2, each group member would receive an overall grade of 85, which would be
fair and appropriate since that was the overall group grade.

**Example 2: Group Grade with Peer Evaluations Factored In**

- Group has 4 people: A, B, C, and D and receive a group grade of an 85
- Each group member rates themselves and their group members on the 3 dimensions. Then the instructor calculates the average for the 3 components for each group member.
- The average ratings for each group member are: A=100, B=100, C=100, and D=100
- 50% of the grade comes from the instructor grade, and 50% is determined from the relative peer member evaluations (as shown below)

<table>
<thead>
<tr>
<th>Overall Instructor Given Group Grade of 85 on the Project</th>
<th>Avg. of the 3 Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Member</td>
<td>Rater</td>
</tr>
<tr>
<td>A</td>
<td>X 100 100 100</td>
</tr>
<tr>
<td>B</td>
<td>100 X 100 100</td>
</tr>
<tr>
<td>C</td>
<td>100 100 X 100</td>
</tr>
<tr>
<td>D</td>
<td>100 100 100 X</td>
</tr>
<tr>
<td>Total</td>
<td>1200</td>
</tr>
</tbody>
</table>

**OUTCOMES AND FEEDBACK FROM STUDENTS**

This team evaluation tool has been utilized four different semesters, in ten different courses ranging from sophomore level classes to capstone senior level courses. In total, this evaluation tool has been completed by 282 students who worked in 78 different class teams (groups). The average team size was 3.62 students per group. Student feedback from this evaluation tool has been primarily, although not exclusively, positive. However, revisions of the evaluation form based on student feedback have resulted in an improved tool that should accurately assess student contributions to teamwork in classes.

In terms of quantitative feedback, we are still assessing the evaluation tool and collecting data to run statistical analyses on a number of factors including perceived fairness, differences in work efforts, how this evaluation tool impacted overall student grades, if this team evaluation tool impacted satisfaction with group projects, and more. However, we did collect data on student perceptions of the different evaluation criteria being “good criteria”, the need for accountability in teamwork, and the overall peer evaluation tool being fair. Students responded to questions on a 5-point Likert type scale with response anchors of 1 = Strongly Disagree and 5 = Strongly Agree. The questions asked, means, and standard deviations are provided in Table 2 below.

**Table 2**

<table>
<thead>
<tr>
<th>Item Questions, Means, and Standard Deviations for Quantitative Evaluation Tool Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question</td>
</tr>
<tr>
<td>1. The peer evaluation process used in this class is fair.</td>
</tr>
<tr>
<td>2. Team member accountability is necessary in teamwork.</td>
</tr>
<tr>
<td>3. The peer evaluation criteria of “quality of work” is a good criteria.</td>
</tr>
<tr>
<td>4. The peer evaluation criteria of “quantity of work” is a good criteria.</td>
</tr>
<tr>
<td>5. The peer evaluation criteria of “being a good team member” is a good criteria.</td>
</tr>
</tbody>
</table>

As can be seen in Table 2, quantitative feedback received from students about the evaluation tool was quite positive. In particular, the students thought the evaluation process was fair, agreed at high levels that there was a need for accountability in teamwork, and thought all three of the evaluation criteria utilized were good criteria. Interestingly, students thought that “quality of work” was slightly the least important criteria, with “quantity of work” being more important, and “being a good team member” being the best criteria. These findings are deserving of additional attention, and bring about the potential idea of weighting the criteria differently, as is discussed in a following section.

In discussing the qualitative feedback received, we will begin with some of the positive responses.

1. **Accountability.** Students have mentioned that the evaluation tool helps to promote accountability and reduce incidents of team members engaging in free-riding and social loafing behavior. Further, students were able to see the additional value of the evaluation tool because they knew that their peers would be evaluated and that the peer evaluation component was worth 50% of the overall grade.

2. **Multiple Evaluation Criteria.** Another aspect of the evaluation form that students enjoyed was that they were evaluated on multiple criteria.
They mentioned that peer evaluation tools often just give an overall rating or focus on overall performance. However, they understand that performance has many different dimensions. Thus, they felt this method of evaluation was a more accurate and fair assessment of everyone’s efforts because they received ratings on quality of work, quantity of work, and being a good team member.

(3) Training in Peer Evaluations. Students also mentioned that they enjoyed the brief training that was provided on completing the evaluation form, as well as being told at the beginning of the semester that they would have to complete the form and what things to keep in mind (i.e., objective facts about each group member’s performance). Students were told to provide objective comments to support their grades. For example, as opposed to saying someone was unreliable or a bad group member, they would report actual behaviors such as “Group Member X didn’t respond to 80% of emails” or “Group Member Y didn’t come to two group meetings and showed up unprepared for another one.”

(4) Understand Instructor Perspective. After discussing the overall goal of the evaluation tool to students, we felt like they had a better appreciation for the reasons behind it. In essence, they understood that just like in the real world, your overall group performance is what it is, regardless of if you were in a group with all social loafers or in a group with highly motivated students. With that being said, we received feedback that they understood that after the evaluations the overall grade should remain the same, but that there should be appropriate differentiation to reflect group member efforts.

As with any evaluation tool, there are potential disadvantages to this method. Students provided feedback on aspects they thought could be improved. In particular, the two items discussed below were mentioned as things that should be considered for revision when using this peer evaluation technique.

(1) Criteria Weighted Equally. One concern that a number of students voiced was that they believed the three criteria (quality of work, quantity of work, and being a good team member) should be weighted differently. When we used the evaluation tool, we had the three criteria weighted equally, but students often mentioned that being a good team member should receive additional weight as it could far outweigh something like quality or quantity of work (i.e., someone could be a strong technical performer, but an absolute pain to work with).

(2) Potential to Inflate Self-Grade by Deflating Peer Grades. Some students mentioned that the potential existed to inflate one’s own grade by deflating other grades. This potential was minimized by requiring as many objective comments as possible and by examining all group member evaluations to determine if there was consensus as opposed to only one person viewing another or everyone negatively. With that being said, the nature of this evaluation tool does create the possibility for students to try to make themselves look better at the expense of others [18].

WAYS TO ADAPT THIS EVALUATION TOOL

As previously mentioned, the evaluation tool described in this study has a number of advantages. One of the greatest of these is that the form can be adapted to the various needs of any instructor. In particular, four adaptations come to mind, with the first of these being to reduce, add to, or change the evaluation criteria. The criteria we used were “Quality of Work,” “Quantity of Work,” and “Being a Good Team Member.” However, based on the project, the nature of the class, or the learning goals of the instructor, the criteria on the form can easily be adapted.

Another way to revise this form based on instructor preferences would be to weight criteria differently. In the classes that this form was used, the three criteria were weighted equally. However, student feedback, both quantitative and qualitative, showed that in certain cases they thought some of the criteria (i.e., “Being a Good Team Member”) were more important than others. Additionally, instructors may have their own preferences for weighting whatever criteria they decide to use.

A third way to adapt the evaluation tool to instructor desires is to weight the overall peer evaluation component differently. In the classrooms that this form was utilized and in the examples above, the relative peer evaluation component was worth 50% of the overall grade. Although we based this weighting on our past experience and on research from performance appraisals in the human resource management literature, it in no way means that other weights cannot be utilized.

A fourth and final method for adapting the evaluation tool described would be to conduct multiple evaluations over the course of the semester. The idea behind this would be that the projects are semester long, so there could be multiple evaluations
throughout the semester to match the actual progress of the class teamwork. A hopeful outcome of multiple evaluations would be that students could have their performance change and it would be noted (e.g., someone performs well at the beginning of the semester and receives credit for it before he or she isn’t as strong a performer at the end of the semester), as opposed to an overall end-of-semester evaluation that is likely to have issues related to the recency bias [21].

CONCLUSION

With the considerable use of teamwork in the classroom environment, there is a need for optimal evaluation tools. Although the extant research has reported a number of different team member evaluation techniques, each has its own drawbacks. Thus, this paper provides details on a teamwork evaluation tool that minimizes a number of negative issues that exist when evaluating team members.

REFERENCES


Appendix 1

Group Member Assessment – APPLICATION PROJECT
Evaluate your group members using the form given below. In completing this form, you will distribute a max of 100 points to each of your group members (including yourself). The point here is not to be generous or be harsh, but to be accurate in your assessment. Except for me, no one will see these forms.

Overall “Quality of Work” on the Application Project (100 points max per person)
Group Member Name: ___________________ Numeric Assessment: __________
____________________ Numeric Assessment: __________
____________________ Numeric Assessment: __________
____________________ Numeric Assessment: __________
Group Member Name (if group size=4): ________
Numeric Assessment: __________
Yourself: ______________________ Numeric Assessment: __________
Reasons for the “Quality of Work” Assessments (i.e., objective comments):

Overall “Quantity of Work” on the Application Project (100 points max per person)
Group Member Name: ___________________ Numeric Assessment: __________
____________________ Numeric Assessment: __________
____________________ Numeric Assessment: __________
____________________ Numeric Assessment: __________
Group Member Name (if group size=4): ________
Numeric Assessment: __________
Yourself: ______________________ Numeric Assessment: __________
Reasons for the “Quantity of Work” Assessments (i.e., objective comments):

Overall “Being a Good Team Member” on the Application Project (100 points max per person)
Group Member Name: ___________________ Numeric Assessment: __________
____________________ Numeric Assessment: __________
____________________ Numeric Assessment: __________
____________________ Numeric Assessment: __________
Group Member Name (if group size=4): ________
Numeric Assessment: __________
Yourself: ______________________ Numeric Assessment: __________
Reasons for the “Being a Good Team Member” Assessments (i.e., objective comments):

Any Additional Comments You’d Like to Provide: