

# *WIP: A Preliminary Investigation of Students as Peer Evaluators of Their Team Members*

Fazel Ranjbar

Department of Engineering and  
Computing Education  
University of Cincinnati  
Cincinnati, OH, USA  
Ranjbarf@mail.uc.edu

Jutshi Agarwal

Department of Engineering  
Education  
University at Buffalo, SUNY  
Buffalo, NY, USA  
jutshiag@buffalo.edu

Elahe Vahidi

Department of Engineering and  
Computing Education  
University of Cincinnati  
Cincinnati, OH, USA  
vahidiee@ucmail.uc.edu

Junqiu Wang

Department of Engineering and  
Computing Education  
University of Cincinnati  
Cincinnati, OH, USA  
[wang5ji@ucmail.uc.edu](mailto:wang5ji@ucmail.uc.edu)

Pk Imbrie

Office of the Senior Vice President  
and Provost  
University of Oklahoma  
Norman, OK, USA  
pk.imbrie@ou.edu

**Abstract**— This work-in-progress research paper focuses on the first-year engineering teams' peer evaluation processes. Teamwork has been identified as a necessary skill for professional engineers and thus an ABET learning outcome. Reliable systematic processes to assess teamwork effectiveness are crucial for improving team outcomes and identifying dysfunctional teams in a classroom setting. Evaluating team effectiveness and identifying dysfunctional teams has been traditionally done through peer evaluations. However, there is a lack of evidence on what/how students' perceptions of effective teammates and functional teams impact peer evaluations. Identifying common indicators and behaviors that students consider when evaluating their peers is a first step to exploring the reliability of peers as evaluators, understanding potential biases, and using peer feedback to investigate dysfunctional teams. This paper aims to take this first step and will explore the following question: “*What evidence do students provide to describe poor team behaviors?*”

The study was conducted at a large, public, urban, Midwestern R1 institution. For the first two semesters of the engineering curriculum, students are required to participate actively in team-based projects as part of engineering design thinking courses. The following procedure was used to answer our research question. Students were first asked to respond to a peer evaluation instrument based on a four-factor (Interdependency, Goal setting, Potency, and Trust) team effectiveness model. Then, students were asked to grade their teammates on a cumulative score of 100 for the team and support their distribution choices with open-ended comments on team behaviors. We started with an analysis of students' comments regarding negative behaviors to support their low rating for each teammate. We selected comments that had a high misalignment between the model-based peer evaluation and corresponding distributed ratings, resulting in a sample of 93 students. Next, we conducted inductive coding on the comments to identify frequently mentioned negative behaviors and see if these behaviors were addressed in the model-based peer evaluation. Eight major themes emerged from students' comments.

**Keywords**—*teamwork, peer evaluation, team effectiveness,*

## I. INTRODUCTION

In today's complex and highly dynamic world, effective teamwork is one of the most important skills for professional success [1]. Teamwork in academic settings can improve creativity in addition to enhancing students' understanding of complex concepts [2], [3], [4]. It also may improve students' engagement, leadership and problem-solving skills, critical thinking, and motivation [5], [6], [7], [8], [9], [10]. As the ability to function on a multidisciplinary team is a critical component of success in different areas, it is also one of ABET's critical outcomes [11]. Developing teamwork skills for engineering students is crucial because it can directly influence their future careers, where employers highly value teamwork skills and working in teams [12], [13].

The effectiveness of a team is defined as its ability to achieve its goals. For this definition, the quality of the output, the level of satisfaction of the team, and how well the members work together to complete tasks are also considered [14]. Communication, collaboration, and coordination are all characteristics of effective teams [1]. Team effectiveness measurement is not easy and depends on several factors. Self-assessments, peer assessments, and objective measurements can all be used to measure team effectiveness [1].

ABET suggests assessing teamwork skills through group projects that evaluate deliverables and teamwork processes, peer assessments of contributions, and communication within the team [11]. Therefore, reliable systematic processes to assess teamwork effectiveness are crucial for improving team outcomes and identifying dysfunctional teams in a classroom setting. One of the methods for evaluating team effectiveness and identifying dysfunctional teams has traditionally been through self and peer evaluations [15], [16], [17]. In addition to assessing team effectiveness, peer evaluation can develop teamwork skills [13]. By evaluating each other, teammates are motivated to demonstrate good teamwork skills and contribute

to the success of the team [18]. Peer evaluation can also make instructors aware of potential team problems and enable them to address the problem appropriately [19], [20].

Considering the importance of peer evaluation in assessing team effectiveness and the dimensions mentioned above, it is important to identify potential biases in student evaluations [21]. Several studies have been done on students' perceptions when evaluating their peers [22], [23], [24]. We know that students' biases when evaluating their peers can reduce the validity of peer evaluation instruments [23], [25]. However, there is scarce evidence on how students' predispositions of team effectiveness and their understanding of peer evaluation criteria impact peer evaluations. A first step to address this gap is to understand the common indicators and behaviors that students give importance to when they are asked to evaluate their teammates. In this study, we explore the following question: *What evidence (or indicators) do students use to describe poor team behaviors?* Identifying common indicators that students give emphasis to can help identify dysfunctional teams and also identify potential biases in students' evaluations in future studies.

## II. METHODS

### A. Educational setting

The study was conducted at a large, public, urban, Midwestern R1 institution. For the first two semesters of the engineering program, students take two 3-credit hour courses. A significant part of the course sequence each semester is the opportunity for students to gain design experience while working as a member of a technical team. Besides participating in team-based projects, students are also required to engage in different fundamental content areas such as design process, ethics, algorithmic thinking, modeling, statistics, statics, and electricity in a flipped-classroom setting. Student teams of three or four members are formed by the teaching team at the start of each semester to attain uniform heterogeneity [26] in factors such as prior experiences, knowledge, and demographics.

### B. Theoretical model and measures

For this study, we base our data on a survey administered at the end of the Fall semester in 2023. In the survey, students are asked to respond to a peer evaluation instrument for each team member. This instrument is based on a four-factor (Interdependency, Goal setting, Potency, and Trust) team effectiveness model [22] ( see Table I). There are three items in each factor. Students are asked to respond to those 12 items on a Likert scale: Strongly agree = 5, agree = 4, Neither agree nor disagree = 3, disagree = 2, and strongly disagree = 1. Henceforth, we refer to this instrument as the model-based scale. In this study, a second section for peer evaluation was added, where students were asked to grade their teammates on a cumulative score of 100% for the team. The responses to this question will be referred to as the distributive grades in the rest of this paper. Lastly, students were asked to support their distributive grades with open-ended comments on positive and negative team behaviors, as noted below. To avoid influencing students' responses and to capture unbiased comments, students were not informed that we would explore possible associations between the two sections.

Open-ended questions:

1- In a couple of sentences, briefly describe the positive teamwork behaviors of each team member that supports the rating you provided (if you didn't observe any positive behaviors, write "exhibited no positive behaviors").

2- In a couple of sentences, briefly describe the negative teamwork behaviors of each team member that supports the rating you provided (if you didn't observe any negative behaviors, write "exhibited no negative behaviors").

Table I Model-based Peer Evaluation Instrument [22]

Construct	Items	Description of each item
Interdependency	IN1	Is collaborating well with the team on all in-class and out of the class assignments
	IN2	Is contributing to the team's success by having a clearly assigned roles
	IN3	Is a reliable team member
Goal Setting	GS1	Is often helping the team develop effective strategies
	GS2	Is articulating individual goals that can be achieved with the help of the team.
	GS3	Is actively helping the team establish goals
Potency	PT1	Is helping the team to build a shared confidence in its ability to successfully work together on course assignments.
	PT2	Is often encouraging each team member to believe in the team's ability to succeed no matter what the task
	PT3	Often making the team feel confident in its ability to resolve disagreements.
Trust	TS1	Is always seen as one who is willing to share important information with teammates.
	TS2	Always appears to be willing to have teammates critically review work performed
	TS3	Is rarely seen checking the work of teammates because of confidence in their approach to solving problems.

### C. Data Collection and Data Analysis

For the purpose of this study, data from the cohort of 1120 engineering students from the fall semester of 2023 were analyzed. The study included a sample size of 1109 students who completed the peer evaluation survey. Each student was asked to evaluate all of their team members, including themselves. This gave us 4036 peer evaluation responses. These responses were all treated as independent for the purpose of this analysis as each collective evaluation was for and by a different individual.

We started with an analysis of students' comments regarding negative behaviors to support their low rating for each teammate. In this work-in-progress, we focus on identifying large discrepancies between the two peer evaluation measures (model-based scale and distributive grades) and examine the open-ended comments to understand the behaviors associated with these discrepancies.

We first standardized the average of the responses on the model-based scale. Then, we identified students whose model-based rating was over 50% but were given extremely low scores (lower than 10 out of the cumulative score of 100) by the same peer. In this way, we screened students with a considerable misalignment between the average model-based score and the distributive grade given by a team member. Using this procedure, we filtered 93 students that fit our criteria.

We performed inductive coding on the comments in order to identify the most common negative behaviors. Our aim was to determine if these behaviors were addressed in the model-based scale. As mentioned earlier, we focused on negative behaviors because identifying negative behaviors may contribute to dysfunctional teams. Additionally, students who gave extremely low scores (in the bottom 10%) to their peers likely viewed their peers as ineffective team members.

After analyzing all 93 open-ended responses, 17 codes emerged. We divided them into eight major themes, which are discussed in the next section.

### III. FINDINGS

Following the selection of 93 students, we conducted an inductive analysis of all negative feedback provided by the students to justify their peers' low ratings. Our findings are presented in Table II, where 17 codes have been classified into eight distinct themes. It should be mentioned that in some comments, more than one code was observed. Furthermore, the table's last column indicates the frequency of each theme in the students' feedback. An example of students' comments related to each theme is provided in Table III.

Table II Themes and Codes

Theme	Code	Frequency
Contribution to team tasks	Lack of contribution	34
	Lack of collaboration	
	Lack of reliability in doing tasks	
Communication Skills	Not responding to messages	28
	Hard to reach	
	Mostly quiet	
Meeting Attendance	Limitation in meeting time	30
	Never shows up to meetings	
	Late for meetings	
Value Add to the Team	Has no novel idea to add	11
	Has weak skills	
	Has no input for team meetings	
Self-Oriented	Rely on themselves	4
	Not a team believer	

No meaningful comments		7
Exhibited no negative behaviors		12
Dropped the course		9

Table III Themes and Examples

Theme	Examples
Contribution to team tasks	<i>"does not contribute much"</i>
	<i>"does not put much effort towards completing the tasks the team assigns him."</i>
	<i>"always on her phone during meetings and rarely contributes to group discussion or progress"</i>
Communication Skills	<i>"Often does not respond to group messages"</i>
	<i>"doesn't respond to ANY attempts to communicate with him."</i>
	<i>"He never answered any texts"</i>
Meeting Attendance	<i>"He is not reliable to showing up to meetings. Being able to stay for long amounts of time."</i>
	<i>"He doesn't always come to meetings"</i>
	<i>"Sometimes does not come to team meetings."</i>
Value Add to the Team	<i>"does not provide a lot of input in group meetings."</i>
	<i>"He is commonly detached from the conversation and will rarely offer his input."</i>
	<i>"does not come up with new ideas to help the project."</i>
Self-Oriented	<i>"often times finds his own personal needs above an urgent project. Not a team player."</i>
	<i>"shuts down easily, so when she is asked to complete the code and she runs into a problem, instead of asking for help she gives up very quickly."</i>
Exhibited no negative behaviors	<i>"Exhibited no negative behaviors"</i>
Dropped the course	<i>" He dropped out without telling us"</i>
No meaningful comments	<i>"I don't know", "No negative and no positive behavior observed"</i>

We defined the emerged themes as below:

**Contribution to Team Tasks:** Students highlighted examples where team members failed to participate in assignments actively, did not fulfill their assigned roles, or demonstrated inconsistency in task completion. In general, this theme includes comments indicating a lack of contribution, collaboration, and reliability in completing group tasks. **Communication Skills:** This theme identifies students who were characterized as quiet in group settings and demonstrated unwillingness or ineffectiveness in communicating within the team. In general, this theme includes situations where students discussed their peers who failed to respond to messages related to the team. It also includes **Meeting Attendance**, which highlights examples of students who were consistently late for

meetings, failed to attend meetings or exhibited limitations in their availability for scheduled meetings. It covers issues related to on-time attendance, reliability, and time management within the team.

**Value Add to the Team:** This category reflects examples where some students failed to offer their input, generate new ideas, or actively contribute to team discussions. Additionally, it included observations of peers with weaknesses in both soft and hard skills, hindering their overall contribution to team success.

**Self-Oriented:** Negative behaviors also pointed to students who prioritized personal needs over team goals or were not fully committed to collaborative efforts. Students described examples where team members exhibited self-centered behavior, lacked belief in the team's objectives, or preferred to work individually rather than as part of the team.

**Dropped the course:** includes students who dropped the course. It reflects instances where team members withdrew from the course or team participation, impacting overall team dynamics and performance.

**No meaningful comments:** Some students either provided no meaningful comments or did not address the question posed.

**Exhibited no negative behaviors:** indicated that some students perceived their peers as not exhibiting any negative behaviors. Despite assigning low ratings, these students did not identify specific negative behavior shown by their team members, suggesting a potential misalignment between their perceptions and the evaluation criteria.

It should be mentioned that the theme of "No value to add to the team" was reflected in comments that some team members did not contribute new ideas or input during group meetings. This lack of engagement can hinder team creativity and problem-solving. It also highlights the importance of actively participating in team discussions and offering unique perspectives. On the other hand, the theme of contribution emphasizes the importance of contributing to completing assignments and activities. In total, the difference between the two themes refers to giving ideas in contrast to doing activities respectively.

Since most of the themes fell under the interdependency factor of the model. We further analyzed the average score for students' responses to the three questions related to the interdependency factor. We expected the average interdependency score to be lower than 50%, and the average scores of the other nine questions, covering the remaining three factors, have resulted in a total average score higher than 50%. But in contrast to our expectation, the selected students had an average rating of 59.5% for interdependency. Additionally, the average scores for the other three factors were as follows: Goal Setting at 60.5%, Potency at 69.1%, and Trust at 70.1%.

#### IV. DISCUSSION AND CONCLUSION

The overall goal of this study was to explore the indicators that students use while assessing their individual team members. The findings of this study, which is the first step of a larger study, will be used to develop a process aiming to identify dysfunctional student teams early in the semester. The main goal is to identify as many indicators as possible from

students' responses to course surveys to identify probable team dysfunctionality.

By analyzing peer evaluation scores and open-ended comments, we aimed to identify potential negative behaviors that may have led the team to be a dysfunctional team.

One important finding of the study was that students frequently provided negative feedback related to their peers' lack of interdependency skills. This suggests that they may not have a clear understanding of the three interdependency questions within the model. Despite giving high scores to these questions, students also gave low rating scores and made negative behavior-related comments. These comments highlight behaviors that could be included in the interdependency questions.

By focusing on the misalignment in the scores based on the model-based scale and distributive grade given by the same team member, we expected identifying some negative behaviors that were not included in the model. However, we found that the described negative behaviors were included in the model. Consequently, the misalignment might be related to discrepancies in the students' evaluation of their peers. Future research is suggested to investigate further the reasons behind such discrepancies and how/why students interpret the model-based items differently from their comments.

It is interesting that some students gave low ratings to their peers but also stated that their peers did not exhibit any negative behaviors in their open-ended comments. This discrepancy requires further investigation to understand the reasons behind students' perceptions. One possible explanation is that other team members demonstrated stronger positive behaviors that led the evaluator to allocate a large portion of the grade to them. To further investigate this explanation, it is necessary to study groups in addition to individuals.

#### V. LIMITATION AND FUTURE WORKS

One major limitation of this study was the size of the sample that we used to gather opinions about observed negative behaviors. In this stage, we only included students with extremely low rating scores compared to high average peer evaluation scores. Our criteria generated only 93 individual ratings when compared to the large data set it was extracted from. This could suggest that only a low number of students represent misaligned ratings. A more refined criteria is necessary to understand the nuances of the student peer evaluations. Furthermore, to gain a more comprehensive understanding of negative behaviors that may affect team functioning, further investigation is necessary, including analyzing both negative and positive comments in addition to analyzing comments at the group level. Analyzing comments at a group level may make it more feasible to identify a dysfunctional team or team member when most or all team members mention a negative behavior observed in one of their team members. Additionally, a comparative analysis of ratings of the same student by different team members will help to identify students' predispositions and other sources of dysfunctionality within teams.

## REFERENCES

- [1] J. E. Mathieu, P. T. Gallagher, M. A. Domingo, and E. A. Klock, "Embracing Complexity: Reviewing the Past Decade of Team Effectiveness Research," *Annu. Rev. Organ. Psychol. Organ. Behav.*, vol. 6, no. 1, pp. 17–46, Jan. 2019, doi: 10.1146/annurev-orgpsych-012218-015106.
- [2] R. Lingard and S. Barkataki, "Teaching teamwork in engineering and computer science," in *2011 Frontiers in Education Conference (FIE)*, Rapid City, SD, USA: IEEE, Oct. 2011, pp. F1C-1-F1C-5. doi: 10.1109/FIE.2011.6143000.
- [3] R. Reiter-Palmon and J. J. Illies, "Leadership and creativity: Understanding leadership from a creative problem-solving perspective," *Leadersh. Q.*, vol. 15, no. 1, pp. 55–77, Feb. 2004, doi: 10.1016/j.leaqua.2003.12.005.
- [4] R. E. Slavin, "Cooperative Learning and Academic Achievement: Why Does Groupwork Work? [Aprendizaje cooperativo y rendimiento académico: ¿por qué funciona el trabajo en grupo?]," *An. Psicol.*, vol. 30, no. 3, pp. 785–791, Oct. 2014, doi: 10.6018/analesps.30.3.201201.
- [5] J. E. Mathieu, M. M. Luciano, L. D'Innocenzo, E. A. Klock, and J. A. LePine, "The Development and Construct Validity of a Team Processes Survey Measure," *Organ. Res. Methods*, vol. 23, no. 3, pp. 399–431, Jul. 2020, doi: 10.1177/1094428119840801.
- [6] Y. Chang and P. Brickman, "When Group Work Doesn't Work: Insights from Students," *CBE—Life Sci. Educ.*, vol. 17, no. 3, p. ar52, Sep. 2018, doi: 10.1187/cbe.17-09-0199.
- [7] J. R. Hackman, "From causes to conditions in group research," *J. Organ. Behav.*, vol. 33, no. 3, pp. 428–444, Apr. 2012, doi: 10.1002/job.1774.
- [8] J. E. Mathieu, J. R. Hollenbeck, D. Van Knippenberg, and D. R. Ilgen, "A century of work teams in the Journal of Applied Psychology," *J. Appl. Psychol.*, vol. 102, no. 3, pp. 452–467, 2017, doi: 10.1037/apl0000128.
- [9] T. Rapp, T. Maynard, M. Domingo, and E. Klock, "Team Emergent States: What Has Emerged in The Literature Over 20 Years," *Small Group Res.*, vol. 52, no. 1, pp. 68–102, Feb. 2021, doi: 10.1177/1046496420956715.
- [10] G. Badea and E. Popescu, "LearnEval Peer Assessment Platform: Iterative Development Process and Evaluation," *IEEE Trans. Learn. Technol.*, vol. 15, no. 3, pp. 421–433, Jun. 2022, doi: 10.1109/TLT.2022.3185423.
- [11] "2022-23-ETAC-Criteria.pdf." Accessed: Apr. 30, 2024. [Online]. Available: <https://www.abet.org/wp-content/uploads/2022/01/2022-23-ETAC-Criteria.pdf>
- [12] E. Salas *et al.*, "Does Team Training Improve Team Performance? A Meta-Analysis," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 50, no. 6, pp. 903–933, Dec. 2008, doi: 10.1518/001872008X375009.
- [13] A. C. Loignon, D. J. Woehr, J. S. Thomas, M. L. Loughry, M. W. Ohland, and D. M. Ferguson, "Facilitating Peer Evaluation in Team Contexts: The Impact of Frame-of-Reference Rater Training," *Acad. Manag. Learn. Educ.*, vol. 16, no. 4, pp. 562–578, Dec. 2017, doi: 10.5465/amle.2016.0163.
- [14] M. A. West, *Effective teamwork: Practical lessons from organizational research*. John Wiley & Sons, 2012.
- [15] M. W. Ohland *et al.*, "The Comprehensive Assessment of Team Member Effectiveness: Development of a Behaviorally Anchored Rating Scale for Self- and Peer Evaluation," *Acad. Manag. Learn. Educ.*, vol. 11, no. 4, pp. 609–630, Dec. 2012, doi: 10.5465/amle.2010.0177.
- [16] D. Anne Delaney, M. Fletcher, C. Cameron, and K. Bodle, "Online self and peer assessment of team work in accounting education," *Account. Res. J.*, vol. 26, no. 3, pp. 222–238, Nov. 2013, doi: 10.1108/ARJ-04-2012-0029.
- [17] M. L. Loughry, M. W. Ohland, and D. J. Woehr, "Assessing Teamwork Skills for Assurance of Learning Using CATME Team Tools," *J. Mark. Educ.*.
- [18] C. Zhou, S. Choi, B. Beigpourian, S. Wei, D. M. Ferguson, and M. W. Ohland, "The Difference between Teams with No Female Students and Teams with Female Students for Peer Evaluation Behavior in Engineering Education," 2019.
- [19] M. Borrego, J. Karlin, L. D. McNair, and K. Beddoes, "Team Effectiveness Theory from Industrial and Organizational Psychology Applied to Engineering Student Project Teams: A Research Review," *J. Eng. Educ.*, vol. 102, no. 4, pp. 472–512, Oct. 2013, doi: 10.1002/je.20023.
- [20] R. Anson and J. A. Goodman, "A Peer Assessment System to Improve Student Team Experiences," *J. Educ. Bus.*, vol. 89, no. 1, pp. 27–34, Jan. 2014, doi: 10.1080/08832323.2012.754735.
- [21] J. B. Main and M. Sanchez-Pena, "Student evaluations of team members: Is there gender bias?," in *2015 IEEE Frontiers in Education Conference (FIE)*, Camino Real El Paso, El Paso, TX, USA: IEEE, Oct. 2015, pp. 1–6. doi: 10.1109/FIE.2015.7344177.
- [22] J. Wang and P. K. Imbrie, "Assessing Team Effectiveness: Comparing Peer Evaluations To A Team Effectiveness Instrument," in *2009 Annual Conference & Exposition Proceedings*, Austin, Texas: ASEE Conferences, Jun. 2009, p. 14.249.1-14.249.9. doi: 10.18260/1-2--5770.
- [23] J. H. Stonewall, "Development and evaluation of training to increase student perceptions of fairness in peer assessment," Iowa State University, 2022.
- [24] F. Ranjbar, J. Agarwal, and P. K. Imbrie, "Work in Progress: Teamwork Predisposition as an Indicator of Team Effectiveness in First-Year Engineering," in *2024 ASEE Annual Conference & Exposition*, 2024.
- [25] J. Wang, "Assessment of engineering student team effectiveness," 2014.
- [26] J. Agarwal, E. Piatt, and P. K. Imbrie, "Team formation in engineering classrooms using multi-criteria optimization with genetic algorithms," in *2022 IEEE Frontiers in Education Conference (FIE)*, Oct. 2022, pp. 1–6. doi: 10.1109/FIE56618.2022.9962741.