The Future of
Engineering Education
2024 Annual Conference & Exposition

Oregon Convention Center
Portland, OR . June 23 - 26, 2024

ASEE

Paper ID #43152

Assessing Student Perceptions of Peer Review Methods' Efficacy in a Team-Based, Senior Undergraduate Capstone Course Setting

Prof. Sara Lego, Pennsylvania State University

Prof. Sara E. Lego joined the Aerospace Engineering Department in 2021 as an Associate Teaching Professor managing the department's four senior undergraduate course offerings. Prior to joining the department, Mrs. Lego spent 20 years within the aerospace industry at Boeing, Iridium, and the Penn State Applied Research Lab as a research engineer, orbital analyst, and engineering program manager. Mrs. Lego received a B.S. in Aerospace Engineering from the Pennsylvania State University in 2001, and an M.S. in Astrodynamics from the University of Colorado at Boulder in 2004. Her research interests include conceptual design, trade space software development, system-of-systems modeling, geospatial situational awareness tools, and visual analytics.

Cara Exten, Pennsylvania State University

Assessing Student Perceptions of Peer Review Methods' Efficacy in a Team-Based, Senior Undergraduate Capstone Course Setting

Abstract

Within Penn State, the Aerospace Engineering department has historically used the Comprehensive Assessment of Team Member Effectiveness (CATME) online tool hosted by Purdue University to conduct its peer reviews within four of its capstone courses, which are year-long offerings running fall to spring. While the tool is research-backed and widely used within capstone courses across the country, recent feedback at Penn State has revealed a perception among students that the tool's "five teamwork dimensions" are overly complex, unclear, and do not allow students to accurately represent true team dynamics and individual behaviors when using the tool to generate peer review scores.

In response to this feedback, the Aerospace Engineering capstone course instructors have developed a new peer review method called Participation and Professional Engineering Skills Assessments (PEPSA), which combines a Participation Factor (PF) approach with a simplified, custom peer review survey generated in Qualtrics that uses a Likert scale and measures the degree to which students agree or disagree with statements related to each team member's performance and professional skills demonstration. This paper describes both the new peer review tool as well as results from a study conducted in the 2022/2023 academic year to evaluate student perceptions of PEPSA against the prior CATME baseline using two identical study questionnaires.

Introduction

Similar to its other peer institutions, Penn State requires its undergraduate engineering students to complete a capstone design project in their senior year. While the duration, topic area, and level of inter-departmental collaboration of the capstone design experience varies across majors, the one feature that is consistent across all capstone offerings is the emphasis on a team-based approach that mimics the industry work environment. While the intent of these capstone courses is to provide students with a means of applying and integrating core concepts within their discipline to a complex problem, the importance of using these courses as a "training ground" for introducing, practicing, and refining team-centric, professional skills that will be critical for students' future career success can not be understated. In fact, the Accreditation Board for Engineering and Technology (ABET) specifically identifies the need to emphasize these "soft" skills within the engineering curriculum as part of its Criterion 3.3, 3.4, and 3.5, which include the following student outcomes [1]:

Criterion 3.3 *An ability to communicate effectively with a range of audiences*

Criterion 3.4 *An ability to recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts*

Criterion 3.5 *An ability to function effectively on a team whose members together provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives*

Given the recognition that development of non-technical skills like communication, task accountability, work quality, professional behavior, and social-emotional intelligence comprises a significant portion of the desired educational outcomes for a graduating engineer, it should come as no surprise that Penn State's team-based capstone course learning objectives map directly to these criteria. It is therefore necessary for instructors and course designers to employ methods to assess individual student outcomes in these areas. However, accurate assessment of individual student performance and contribution to overall team productivity can be difficult in capstones, where 50% or more of the grade is often determined based on team-generated content. Significant research has been conducted on methods and techniques that can be used to inform teacher evaluations of individual professional skills and their impact on team dynamics, with weekly team reports, meeting notes, team "time cards", professor-led peer feedback sessions, self-identified contribution scores, and survey-based peer reviews all being suggested as possible options [2]-[8]. Of these techniques, peer reviews have consistently demonstrated their efficacy as a tool for professors to gain insight into the contributions, participation, and professional skills development of individual team members, with combined peer review results from all team members offering a further snapshot of positive and negative team dynamics [9]-[12]. Unfortunately, the way in which peer review methods are implemented (anonymous vs. identifiable, digital vs. in-person), as well as question content and clarity, have been shown to drastically impact the degree to which peer review results accurately reflect the actual interactions of the team and its members' behaviors [10]-[13]. Furthermore, if students perceive a peer review tool as non-representative and inadequate in its ability to match their perceptions of internal team dynamics, then those perceptions can impact student value assessments of the capstone experience as a whole. A sense of "grade injustice" in the presence of social loafing, poor quality work, or communication/interpersonal deficiencies can be demotivating for otherwise high performing students if they reject the adequacy of the mechanisms implemented to catch, correct, and resolve these behaviors [14].

Background and study motivation

Within the Aerospace Engineering senior capstone courses, instructors have historically used the Comprehensive Assessment of Team Member Effectiveness (CATME) peer review tool. CATME is a widely used, evidence-based survey tool that allows students to perform self- and peer-evaluations anonymously by rating themselves and team members in 5 core areas: Contributing to Team's Work; Interacting with Teamwork; Keeping the Team on Track; Expecting Quality; and Having Related Knowledge, Skills, and Abilities. Originally developed by Loughry and Ohland et. al through National Science Foundation funding and extended by a diverse team of engineering educators and consultants, the tool contains a number of valuable features, all supported by multiple benchmarked studies, that make it an ideal choice for capstone peer reviews [15]-[17]. The web-based version of the tool is currently hosted by Purdue University, and its user interface, survey setup, distribution capabilities, and documentation are both user-friendly and straightforward. Following initial class roster setup, CATME notifies students that a survey activity is open, will remind students prior to activity close, and automates the process of aggregating scoring and collating anonymous comments by generating personalized reports that the tool itself will send to each survey participant. Unlike some peer assessment methods that leave student ratings of team skills and behaviors open to interpretation by the student, CATME includes a Behaviorally Anchored Ratings Scale (BARS) instrument in

which each scoring level is mapped to specific, demonstrable behaviors and characteristics [16]. The ratings scale-behavior mapping (Figure 1) is available to students at the instructor's discretion and is easily accessible online [18].

Score	Contributing to Team's Work	Interacting with Teammates	Keeping the Team on Track	Expecting Quality	Having Related Knowledge, Skills, and Abilities
5	• Does more or higher-quality work than expected. • Makes important contributions that improve the team's work. • Helps teammates who are having difficulty completing their work.	• Asks for and shows an interest in teammates' ideas and contributions. • Makes sure teammates stay informed and understand each other. • Provides encouragement or enthusiasm to the team. • Asks teammates for feedback and uses their suggestions to improve.	• Watches conditions affecting the team and monitors the team's progress. • Makes sure that teammates are making appropriate progress. • Gives teammates specific, timely, and constructive feedback.	• Motivates the team to do excellent work. • Cares that the team does outstanding work, even if there is no additional reward. • Believes that the team can do excellent work.	• Demonstrates the knowledge, skills, and abilities to do excellent work. • Acquires new knowledge or skills to improve the team's performance. • Able to perform the role of any team member if necessary.
4	Demonstrates behaviors described immediately above and below.				
3	• Completes a fair share of the team's work with acceptable quality. • Keeps commitments and completes assignments on time. • Helps teammates who are having difficulty when it is easy or important.	• Listens to teammates and respects their contributions. • Communicates clearly. Shares information with teammates. • Participates fully in team activities. • Respects and responds to feedback from teammates.	• Notices changes that influence the team's success. • Knows what everyone on the team should be doing and notices problems. • Alerts teammates or suggests solutions when the team's success is threatened.	• Encourages the team to do good work that meets all requirements. • Wants the team to perform well enough to earn all available rewards. • Believes that the team can fully meet its responsibilities.	• Demonstrates sufficient knowledge, skills, and abilities to contribute to the team's work. • Acquires knowledge or skills as needed to meet requirements. • Able to perform some of the tasks normally done by other team members.
2	Demonstrates behaviors described immediately above and below.				
1	• Does not do a fair share of the team's work. Delivers sloppy or incomplete work. • Misses deadlines. Is late, unprepared, or absent for team meetings. • Does not assist teammates. Quits if the work becomes difficult.	• Interrupts, ignores, bosses, or makes fun of teammates. • Takes actions that affect teammates without their input. Does not share information. • Complains, makes excuses, or does not interact with teammates. • Is defensive. Will not accept help or advice from teammates.	• Is unaware of whether the team is meeting its goals. • Does not pay attention to teammates' progress. • Avoids discussing team problems, even when they are obvious.	• Satisfied even if the team does not meet assigned standards. • Wants the team to avoid work, even if it hurts the team. • Doubts that the team can meet its requirements.	• Missing basic qualifications needed to be a member of the team. • Unable or unwilling to develop knowledge or skills to contribute to the team. • Unable to perform any of the duties of other team members.

Figure 1: CATME Teamwork Rating Scale and Associated Behaviors for Each Category [18]

To ensure consistency across ratings, the CATME developers have also incorporated a training feature that allows students to assess a set of imaginary teammates given some suggested behavior characteristics against the BARS instrument. The intent is to calibrate student ratings to align more closely with the 5 point Likert scale shown in Figure 1 instead of students using an arbitrary interpretation of these levels. Given all of these features, the accessibility of the tool via its online interface, and its research foundation, it's easy to see why the tool enjoys widespread use within undergraduate capstone courses throughout the United States.

Despite these advantages, feedback from 2020-2022 Penn State Aerospace Engineering capstones has revealed a pervasive perception among the undergraduate 4th-year students that CATME's "five teamwork dimensions" are overly complex, unclear, and do not allow students to accurately represent perceived team dynamics. They also expressed concerns over the use of the tool for assigning a peer review grade within the course since the 5-point Likert scale CATME uses doesn't map readily to a 100 percentage point grading scheme. As a result, students indicated that they approached the peer review scoring process with some trepidation, as a "3" score in all categories might result in a teammate receiving a "C" peer review grade in the course. These criticisms came as a surprise to capstone instructors, as the tool had been successfully used within the department for over a decade. In addition, the CATME training feature has always been mandated for the first peer review survey of each capstone course, so the opportunity for students to calibrate their ratings against the CATME behavior mapping matrix should not have changed. Upon further discussion with students, however, it became clear that

the students' interpretation of the rating scale and the behavior anchors wasn't consistent with prior cohorts. For example, the 2020-2022 classes voiced confusion over why level 3 behaviors shown in Figure 1 were less desirable than level 5 ones, and felt that the differences between them were too nuanced . It appeared possible that students' perceptions of what behaviors they valued most in a team was shifting. While the value evolution of student teamwork traits is an intriguing one, a more immediate concern for the capstone courses was how to effectively assess individual contribution while motivating development of professional skills in a clear way that correlates with student perception of team dynamics. Therefore, a new peer review tool was developed to address these concerns, and a study was conducted to compare student perceptions of the two methods.

Peer review method development and comparison study

The discrepancy between the CATME rating scale's behavior anchors and student valuation of teamwork traits seemed to be negatively impacting student perceptions of the peer review method's ability to accurately reflect team dynamics. The Aerospace Engineering capstone instructors consequently developed a new peer review method called the Participation and Professional Engineering Skills Assessment (PEPSA), which combines the Participation Factor (PF) approach identified by Friess and Goupee [10] with a simplified, custom peer review survey generated in Qualtrics.

Building upon the CATME development research from Loughry and Ohland et al [16]-[17], the survey includes 20 statements that reflect the behavioral traits most representative of teamwork success as well as questions related to fostering an environment of psychological safety, inclusivity, and respect. While these "team climate" types of questions are available within CATME as additional peer review survey options, the tool uses them to assess each respondent's personal feeling of safety within the group. In contrast, the PEPSA survey uses these questions as peer rating opportunities in these areas. In effect, the PEPSA survey removes the top layer of the CATME interface in which behaviors are combined into categories, instead presenting students with clear, single behavior statements related to each team member's performance and professional skills demonstration. The PEPSA survey then asks students to rate the degree to which a teammate demonstrated a skill or behavior on a 10 point, anchored Likert scale for each of the 20 statements. The 10-point Likert scale was chosen over the 5-point CATME scale to map more closely with a 100% point grading scheme and address student concerns related to peer review grade assignments.

In addition to the custom survey, which is administered 2-3 times during the course of each semester, the PEPSA method includes a short, two question Teammate Contribution Factor (TCF) survey following each assignment that asks students to allocate a percentage grade out of 100% in two critical areas for that assignment only:
1. Communication quality and frequency/responsiveness (50%)
2. Work quality and completion (50%).

This approach adjusts Friess and Goupee's [10] weekly PF peer evaluation by tying it to a specific assignment while also adding a communication metric and percentage score requirement. The resulting TCF for each teammate is then calculated using the same PF equation (Eq. 1) proposed in [10]:

$$TCF = \frac{Average\ individual\ TCF\ grade}{Average\ TCF\ grades\ of\ all\ team\ members} \qquad \text{(Eq. 1)}$$

The resultant individual TCF was used to adjust the team assignment grade for each student.

Since Penn State Aerospace Engineering capstone courses are year-long offerings, they provided an ideal environment in which to pilot PEPSA for a single semester and study student perceptions of the tool's efficacy as a peer review method compared to CATME. Four capstone courses comprising 117 students used CATME for 3 peer reviews during the Fall 2022 semester, after which they were offered an anonymous Qualtrics study questionnaire about the tool. During the Spring 2023 semester, the PEPSA tool, including 3 peer review surveys and per-assignment TCF surveys, was piloted. An identical, anonymous Qualtrics study questionnaire about PEPSA was then offered to the same capstone cohort (reduced to 114 students due to December graduations) at the end of the 2022-2023AY.
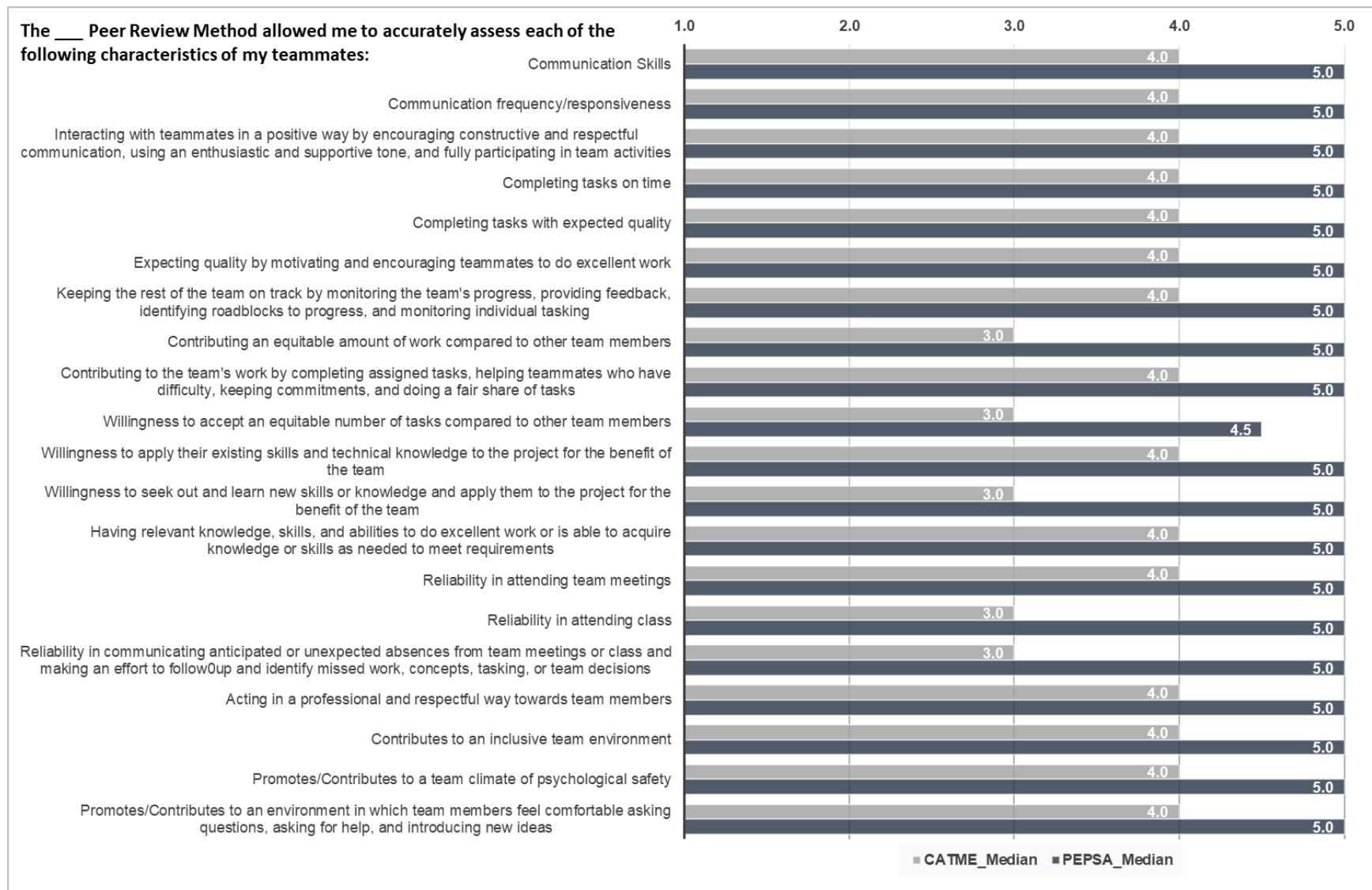
The questionnaire requested students to determine the degree to which each peer review method allowed them to accurately assess individual characteristics and skills that contribute to a positive team experience. In addition, students were asked to provide an anchored, 5-point Likert scale rating for 6 statements about the peer review method's administration within the class and its ability to reflect their perception of team dynamics. To meet human study IRB requirements, no incentive was offered for completion of the surveys since the study author is an instructor for the capstone courses.

Study results and interpretation

Survey response rates were low, likely due to the lack of incentives offered, with a 24.8% response rate for the CATME questionnaire and a 15.8% response rate for the PEPSA questionnaire. Since the study was conducted anonymously, there was no pairing between response data for the two questionnaires.

Figures 2 and 3 compare student median Likert score values for the CATME and PEPSA tools broken down by the two question categories contained within the study questionnaires. Figure 2 provides median values per behavior/skills trait statements, while Figure 3 shows median scores specifically related to questions pertaining to the peer review method's administration and student perception of its efficacy as a peer review tool. For both question sets, PEPSA median scores surpassed CATME by at least 1 full point except in relation to administration frequency (median = 4.5 for PEPSA vs. 4.0 for CATME). PEPSA response variance was also lower, sometimes by as much as 50%, for 92.3% of questions, indicating higher scoring agreement among PEPSA respondents.

Next, the authors compared student perception scores between CATME and PEPSA. The mean, median, and variance were calculated for each statement on the questionnaire. Because of the small sample size, assumptions for a t-test were likely not met. Therefore, the Mann-Whitney test was used to compare the median evaluation scores associated with the CATME and PEPSA tools. Results from the Mann-Whitney tests reveal that student perceptions of the CATME and PEPSA differed significantly (p-value <0.05) on every item assessed (Tables 1 and 2).

The ___ Peer Review Method allowed me to accurately assess each of the following characteristics of my teammates:

Characteristic	CATME_Median	PEPSA_Median
Communication Skills	4.0	5.0
Communication frequency/responsiveness	4.0	5.0
Interacting with teammates in a positive way by encouraging constructive and respectful communication, using an enthusiastic and supportive tone, and fully participating in team activities	4.0	5.0
Completing tasks on time	4.0	5.0
Completing tasks with expected quality	4.0	5.0
Expecting quality by motivating and encouraging teammates to do excellent work	4.0	5.0
Keeping the rest of the team on track by monitoring the team's progress, providing feedback, identifying roadblocks to progress, and monitoring individual tasking	4.0	5.0
Contributing an equitable amount of work compared to other team members	3.0	5.0
Contributing to the team's work by completing assigned tasks, helping teammates who have difficulty, keeping commitments, and doing a fair share of tasks	4.0	5.0
Willingness to accept an equitable number of tasks compared to other team members	3.0	4.5
Willingness to apply their existing skills and technical knowledge to the project for the benefit of the team	4.0	5.0
Willingness to seek out and learn new skills or knowledge and apply them to the project for the benefit of the team	3.0	5.0
Having relevant knowledge, skills, and abilities to do excellent work or is able to acquire knowledge or skills as needed to meet requirements	4.0	5.0
Reliability in attending team meetings	4.0	5.0
Reliability in attending class	3.0	5.0
Reliability in communicating anticipated or unexpected absences from team meetings or class and making an effort to follow0up and identify missed work, concepts, tasking, or team decisions	3.0	5.0
Acting in a professional and respectful way towards team members	4.0	5.0
Contributes to an inclusive team environment	4.0	5.0
Promotes/Contributes to a team climate of psychological safety	4.0	5.0
Promotes/Contributes to an environment in which team members feel comfortable asking questions, asking for help, and introducing new ideas	4.0	5.0

Figure 2: Peer Review Method (CATME vs. PEPSA) Median Likert Score Rating for 20 Individual Behavior and Professional Skills Assessment Metrics

Table 1: CATME vs. PEPSA Peer Review Method Study Results - Individual Behavior Traits/Skills Question Set

"The ___ Peer Review Method allowed me to accurately assess each of the following characteristics of my teammates:"	CATME Peer Review Method n=29, σ=0.3			PEPSA Peer Review Method n=18, σ=0.2			Mann-Whitney p-value
	Mean	Median	Variance	Mean	Median	Variance	
Communication Skills	3.5	4.0	0.7	4.7	5.0	0.3	<.0001
Communication frequency/responsiveness	3.2	4.0	0.9	4.6	5.0	0.3	<.0001
Interacting with teammates in a positive way by encouraging constructive and respectful communication, using an enthusiastic and supportive tone, and fully participating in team activities	3.5	4.0	0.7	4.7	5.0	0.2	<.0001
Completing tasks on time	3.3	4.0	1.3	4.8	5.0	0.1	<.0001
Completing tasks with expected quality	3.8	4.0	0.9	4.7	5.0	0.2	0.0007
Expecting quality by motivating and encouraging teammates to do excellent work	3.6	4.0	0.8	4.4	5.0	1.0	0.0013
Keeping the rest of the team on track by monitoring the team's progress, providing feedback, identifying roadblocks to progress, and monitoring individual tasking	3.2	4.0	1.5	4.5	5.0	0.7	0.0004
Contributing an equitable amount of work compared to other team members	3.0	3.0	1.1	4.6	5.0	0.3	<.0001
Contributing to the team's work by completing assigned tasks, helping teammates who have difficulty, keeping commitments, and doing a fair share of tasks	3.3	4.0	0.8	4.5	5.0	0.7	0.0002
Willingness to accept an equitable number of tasks compared to other team members	3.0	3.0	1.1	4.2	4.5	1.0	0.0007
Willingness to apply their existing skills and technical knowledge to the project for the benefit of the team	3.6	4.0	1.0	4.6	5.0	0.5	0.0003
Willingness to seek out and learn new skills or knowledge and apply them to the project for the benefit of the team	3.1	3.0	1.2	4.7	5.0	0.3	<.0001
Having relevant knowledge, skills, and abilities to do excellent work or is able to acquire knowledge or skills as needed to meet requirements	3.5	4.0	1.2	4.5	5.0	0.4	0.0011
Reliability in attending team meetings	3.3	4.0	1.8	4.6	5.0	0.7	0.001
Reliability in attending class	2.8	3.0	1.9	4.4	5.0	1.5	0.0002
Reliability in communicating anticipated or unexpected absences from team meetings or class and making an effort to follow0up and identify missed work, concepts, tasking, or team decisions	3.0	3.0	2.0	4.6	5.0	0.7	0.0002
Acting in a professional and respectful way towards team members	3.6	4.0	1.4	4.7	5.0	0.3	0.0003
Contributes to an inclusive team environment	3.7	4.0	1.5	4.6	5.0	0.5	0.0067
Promotes/Contributes to a team climate of psychological safety	3.4	4.0	1.8	4.5	5.0	0.5	0.0052
Promotes/Contributes to an environment in which team members feel comfortable asking questions, asking for help, and introducing new ideas	3.8	4.0	1.1	4.5	5.0	0.4	0.0107
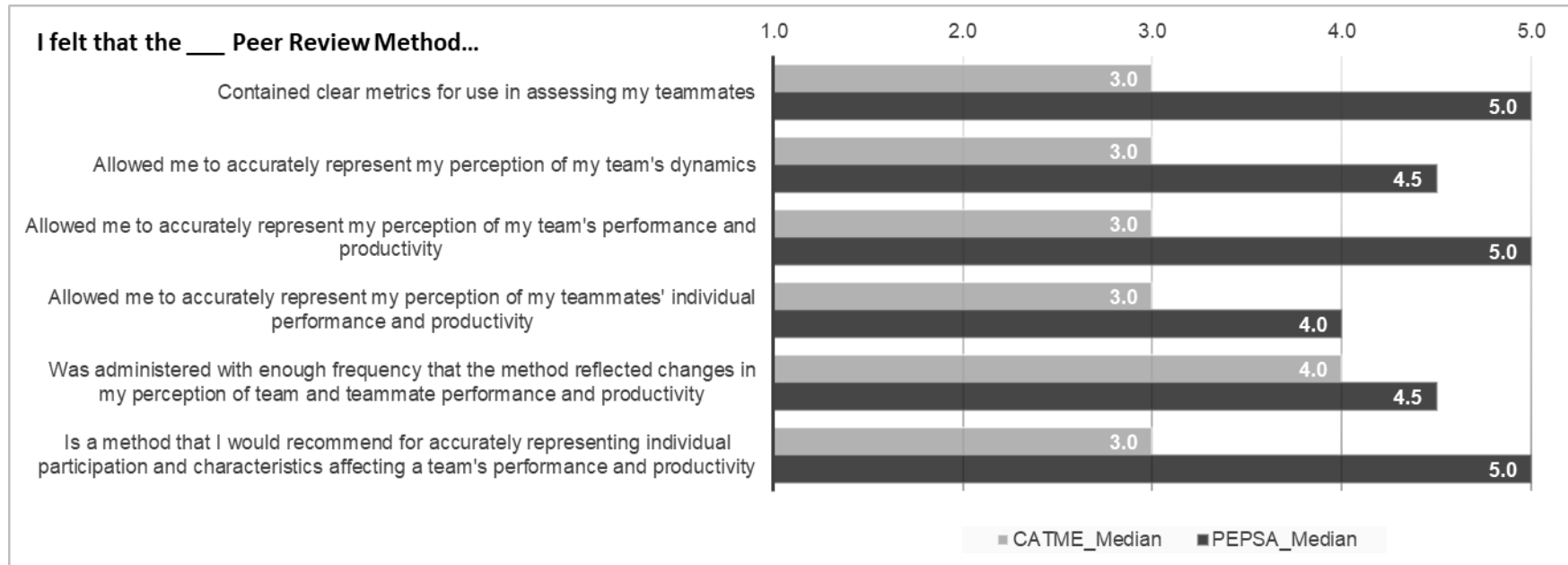
Figure 3: Peer Review Method (CATME vs. PEPSA) Median Likert Score Rating for Method Perception Questions

Table 2: CATME vs. PEPSA Peer Review Method Study Results - Student Perception of Tool Efficacy Question Set

"I felt that the ___ Peer Review Method..."	Mean	Median	Variance	Mean	Median	Variance	Mann-Whitney p-value
Contained clear metrics for use in assessing my teammates	2.8	3.0	1.9	4.7	5.0	0.3	<.0001
Allowed me to accurately represent my perception of my team's dynamics	2.8	3.0	1.8	4.2	4.5	1.0	0.0006
Allowed me to accurately represent my perception of my team's performance and productivity	3.0	3.0	1.3	4.5	5.0	0.4	<.0001
Allowed me to accurately represent my perception of my teammates' individual performance and productivity	3.1	3.0	2.1	4.3	4.0	0.6	0.0039
Was administered with enough frequency that the method reflected changes in my perception of team and teammate performance and productivity	3.3	4.0	1.7	4.1	4.5	1.7	0.0328
Is a method that I would recommend for accurately representing individual participation and characteristics affecting a team's performance and productivity	2.8	3.0	1.6	4.3	5.0	1.0	0.0004

In sum, findings suggest statistically significant differences between student perceptions of CATME and PEPSA's ability to accurately reflect team dynamics and teammate professional behavior, inclusivity, and skills traits.

Discussion

While these initial results are positive, the authors acknowledge that they should be interpreted with some caveats. Sample sizes for both the CATME and PEPSA study questionnaires were small, and no claims are made related to the PEPSA tool's ability to accurately reflect or predict actual teammate behaviors since an alternate, established peer review method was not included as a control. The study instead aimed to assess each tool's ability to reflect student *perceptions* of team dynamics and teammate traits. Peer review efficacy *perception* is an individual and potentially fluid assessment based on each person's valuation of teamwork traits and skills. Furthermore, while the PEPSA tool's removal of CATME's trait categorization into 5 core areas may have increased student clarity about what they were assessing (and thus resulting in higher perceived efficacy), it's also possible that a number of other factors resulted in higher median scores for PEPSA over CATME. Student interpretation of a survey tool's effectiveness may be intrinsically linked to implementation of the CATME tool within the course, including grading. Since students are used to thinking within a 100% point system, it's possible that PEPSA's 10-point Likert scale provided more clarity to their score-to-grade internal mapping than the 5-point scale used by CATME. The CATME method was also administered in the first semester of the course, while teams are still forming and students are gradually acclimating to developing team climates. The sequential nature of the study design (CATME administered in Fall and PEPSA in the Spring) made this influence difficult to isolate. Finally, it should be noted that the study questionnaire's behavior and professional skills question set shares phrasing with the PEPSA surveys, which are in turn adaptations of the CATME score matrix. This shared language may have skewed results in favor of the PEPSA tool.

Observations and next steps

Anecdotally, the PEPSA tool was more effort for instructors to setup, maintain, and administer than the CATME tool since it lacks a custom online user interface and is primarily built upon a combination of Qualtrics and several Excel macros for averaging and collating the data. However, students did seem to feel that PEPSA provided a "fairer" grade distribution across team members based on the traits assessed. While it was expected that the inclusion of the per assignment TCF within the PEPSA tool would motivate students and diminish social loafing, there was no clear indication that this occurred. Students who demonstrated poor attendance at team meetings and low work quality/accountability in the Fall semester continued to have these issues regardless of the peer review method employed. While the PEPSA TCF should have given students an ongoing opportunity throughout the Spring semester to capture this negative teamwork behavior in their percentage scores, students were reticent to use the feature as intended. For all tools, there was a tendency for students to numerically score each other highly for fear of impacting their peers' grades or causing team conflict. True team dynamics were then only reflected in the anonymous comments submitted by respondents for their teammates in all peer review methods. Perhaps because of this disconnect, student feedback indicated that they

disliked the TCF portion of the PEPSA tool, citing that it was considerable additional work to rate all team members after each assignment, and resulted in minimal impact on team dynamics.

Based on the study data and lessons learned, capstone instructors retained the PEPSA survey tool for the 2023-2024 AY, but removed the TCF portion. In addition, a "no cost" feature was added to the tool's administration in the Spring 2024 semester such that students could complete the PEPSA survey's scoring and anonymous comment portions without fear of impacting their peers' grades. Instead, the "no cost" surveys are intended to act as a performance review improvement plan without penalty for each student. The surveys are administered during the 1st and 11th week of the 16-week Spring semester, with graded PEPSA surveys conducted during weeks 6 and 16. Future data collection on these adjustments, to include a larger sample size, modified study questionnaire, and parallel administration of the CATME and PEPSA tools in the same semester, are proposed next steps beyond the current study.

References

[1]   ABET, "Criteria for Accrediting Engineering Programs, 2022-2023,". https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2022-2023/ (retrieved December 1, 2022).

[2]   E. Doerry, and J. D. Palmer, "Improving Efficacy of Peer-Evaluation in Team Project Scenarios," in *Proceedings of the 2011 ASEE Annual Conference & Exposition*, Vancouver, BC., June, 2011.

[3]   B. Beigpourian, "Understanding the Relationship Between Team Dynamics on Peer Evaluations and Team Effectiveness," PhD Dissertation, Purdue University, West Lafayette, Indiana, August 2020.

[4]   B.A. Friedman, P.L. Cox, L.E. Maher, "An expectancy theory motivation approach to peer assessment," *Journal of Management Education*, vol. 32, no. 5, pp. 580-612, 2008.

[5]   , M.F. Couturier and G. Bendrich, "Evaluation of individual members in engineering design teams," in *Proceedings of the CEEA Canadian Engineering Education Conference*, CEEA15, Hamilton, ON, 2013.

[6]   M.R. Sandlin, M.R. Price, and K. Perez, "A Capstone Experience: Impacts of a Behavioral Style Learning Unit on Soft Skill Development and Team Dynamics," *Journal of Agricultural Education*, 59(1), 21-34, 2018.

[7]   A. Hurst, and O. Nespoli, "Peer Review in Capstone Design Courses: An Implementation Using Progress Update Meetings," *The International Journal of Engineering Education*, vol. 31, no. 6, 1799-1809, 2015.

[8]   K. Demoret, "Team Time Cards: a Tool to Increase Accountability and Reduce Social Loafing in Senior Design," in *Proceedings of the 2019 ASEE Annual Conference and Exposition,* June, 2019.

[9] S. Li, T. O'Neill, R. Brennan, G. Gress, A.J. Ebufegha, and M. Lee, "Informing Team Dynamics of Capstone Projects Using Peer Evaluation Scores," in *Proceedings of the 2020 Canadian Engineering Education Association (CEEA-ACEG20) Conference*, Paper 052, June, 2020.

[10] W.A. Friess and A.J. Goupee, "Using Continuous Peer Evaluation in Team-Based Engineering Capstone Projects: A Case Study," *IEEE Transactions on Education*, vol. 63, no. 2, pp. 82-87, May, 2020.

[11] Y. Cinar and A. Biglin, "Peer assessment for undergraduate teamwork projects in petroleum engineering," *International Journal of Engineering Education*, vol. 27, no. 2, pp. 310–322, 2011.

[12] N. Falchikov, "Involving students in assessment," *Psychology Learning and Teaching*, vol. 3, no. 2, pp102–108, 2004.

[13] C. Marcangelo, P. Cartney, and C. Barnes, "The opportunities and challenges of self, peer and group assessment," *Contemporary issues in assessment in health sciences and practice education*, pp. 1–20. London, UK: Higher Education Academy. 2010.

[14] U. Samarakoon, A. Imbulpitiya, and K. Manathunga, "Say No to Free Riding: Student Perspective on Mechanisms to Reduce Social Loafing in Group Projects," *Proceedings of the 13th International Conference on Computer Supported Education (CSEDU 2021),* Vol. 1, pp. 198-206, 2021.

[15] M.L. Loughry, M. W. Ohland, and D.D. Moore, "Development of a theory-based assessment of team member effectiveness," *Educational and Psychological Measurement*, vol. 67, pp. 505-524, 2007.

[16] M.W. Ohland, M. L. Loughry, D.J. Woehr, L.G. Bullard, R.M. Felder, C.J. Finelli, R.A. Layton, H.R. Pomeranz, and D.G. Schmucker, "The comprehensive assessment of team member effectiveness: Development of a behaviorally anchored rating scale for self and peer evaluation," *Academy of Management Learning & Education*, vol. 11, no. 4, pp. 609-630, 2012.

[17] A.C. Loignon, D.J. Woehr, J.S. Thomas, L.M. Loughry, M.W. Ohland, and D. Ferguson, "Facilitating Peer Evaluation in Team Contexts: The Impact of Frame-of-Reference Rater Training," *Academy of Management Learning & Education*, vol. 16, no. 4, pp.562-578, 2017.

[18] "CATME Rating Scale." Purdue University, 2024. https://info.catme.org/download/catme-rating-scale/?wpdmdl=3181&masterkey=5f36d3e01d447 (accessed: Dec. 1, 2022).