2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #52952

The Construct Validity Evidence of the Team Development Peer Evaluation Scale in An Engineering Foundation Course

Mr. Fazel Ranjbar, University of Cincinnati

Fazel Ranjbar is a doctoral candidate in Engineering and Computing Education at the University of Cincinnati and holds a Graduate Certificate in Data Analytics. He combines a background in industrial engineering and over a decade of management experience with advanced statistical methods to evaluate educational programs. His academic research investigates team dynamics and teamwork in undergraduate students. To support this work, he builds automated scoring platforms and applies natural language processing to validate educational assessments. His research aims to inform institutional decisions and improve instructional interventions in higher education.

Dr. So Yoon Yoon, University of Cincinnati

Dr. So Yoon Yoon is an assistant professor in the Department of Engineering and Computing Education in the College of Engineering and Applied Science at the University of Cincinnati, OH, USA. Dr. Yoon received her Ph.D. in Gifted Education and an M.S.Ed. in Research Methods and Measurement with a specialization in Educational Psychology, both from Purdue University, IN, USA. She also holds an M.S. in Astronomy and Astrophysics and a B.S. in Astronomy and Meteorology from Kyungpook National University, South Korea. Her work centers on elementary, secondary, and postsecondary engineering education research as a psychometrician, data analyst, and program evaluator with research interests in spatial ability, STEAM education, and research synthesis with a particular focus on meta-analysis. She has developed, validated, revised, and copyrighted several instruments beneficial for STEM education research and practice. Dr. Yoon has authored more than 95 peer-reviewed journal articles and conference proceedings and served as a journal reviewer in engineering education, STEM education, and educational psychology. She has also served as a PI, co-PI, advisory board member, or external evaluator on several NSF- and NASA-funded projects.

Dr. Junqiu Wang, University of North Carolina at Charlotte

Dr Wang currently serves as Teaching Associate Professor at Office of Student Development and Success, William States Lee College of Engineering, University of North Carolina Charlotte. He received his Ph.D. in Engineering Education from Purdue University, M.S. in Materials Engineering from Tsinghua University, China and B.S. in Materials Engineering from University of Sci. & Tech. Beijing, China

Dr. P.K. Imbrie, University of Cincinnati

P.K. Imbrie is the Head and Professor of the Department of Engineering Education and a Professor in the Department of Aerospace Engineering and Engineering Mechanics University of Cincinnati. He received his B.S., M.S. and Ph.D. degrees in Aerospace Engi

The Construct Validity Evidence of the Team Development Peer Evaluation Scale in an Engineering Foundation Course

Abstract

This empirical full paper examines construct validity and reliability evidence for a peer evaluation scale built on a four-factor team effectiveness model. Developing teamwork skills is a critical outcome in engineering programs. However, assessing teamwork skills remains a challenge for educators. Many existing teamwork assessment instruments are either expensive or lack validation for diverse student populations. This study examines the construct validity and reliability of a four-factor peer evaluation instrument, Team Development Peer Evaluation Scale (TDPES), measuring Interdependency, Goal Setting, Potency, and Trust. Data were collected from 2,077 first-year engineering students at a large public university in the Midwestern United States. A confirmatory factor analysis (CFA) was conducted to test the fit of the hypothesized four-factor model for the TDPES. The results indicated a model fit (χ^2 = 651.93, CFI = .977, TLI = .970, RMSEA = .076) and high internal consistency across all constructs with Cronbach's α ranging from .86 to .92 for the four constructs and .96 for the overall scale. These findings support the construct validity of the TDPES and its reliability as one component of a broader construct validity argument. Engineering educators can use this accessible instrument to diagnose team functioning and support the development of teamwork skills, while recognizing that additional validity evidence, such as relations with external outcomes, remains a direction for future work.

1. Introduction

The ability to function effectively on teams is a pillar of modern engineering practice [1]. Professional organizations and accrediting boards like the Accreditation Board for Engineering and Technology (ABET), the Active Learning in Engineering Education (ALE) network, and European Society for Engineering Education (SEFI), have long recognized that the solutions to society's most pressing challenges require collaborative effort that transcends individual contribution and recommended that engineering education programs incorporate active learning and teamwork strategies into their curriculum [1, 2]. Consequently, engineering education programs are directed to cultivate this competency, embedding team-based projects throughout the curriculum to provide students with authentic, hands-on collaborative experiences [1, 3-5].

Over the past two decades, researchers and practitioners have developed a range of measurements and frameworks to support team formation, peer assessment, and teamwork development in engineering education and related fields. Platforms, such as CATME [6], ITP Metrics [7], Teammates [8], and Teams+ [9], provide structured approaches for capturing peer feedback that consequently result in improving team dynamics. Imbrie, et al. [10] developed a four-factor model of team effectiveness to measure student team effectiveness working in a technical engineering team in the first year of engineering programs. Based on the team effectiveness model, the Team Development Peer Evaluation Scale (TDPES), a peer evaluation instrument, was developed to assess engineering students' team effectiveness as perceived by their peers [11]. The instrument has been widely used in several large U.S. universities.

1.1 Purpose of the Study

This research aims to confirm the validity and reliability of the TDPES by answering the following research questions: (a) To what extent does the collected student peer evaluation data fit the hypothesized factor structure of the instrument?; and (b) To what extent does the peer evaluation instrument demonstrate internal consistency reliability?

2. Literature Review

2.1 Team and Teamwork

A team is defined as two or more interdependent individuals working together toward common, shared goals [4, 12-15]. Although all teams are groups, all groups cannot always function as teams [16]. In a team, team members undertake different roles, responsibilities, and interact socially, whether face-to-face or virtually [17, 18]. Teams maintain and manage clear boundaries that distinguish members from non-members [17, 19-21]. Teams with a higher level of coordination, collaboration, and accountability are more likely to achieve shared and common goals [21].

Teamwork is recognized as a vital skill in engineering and other STEM disciplines, central to both professional practice and ABET accreditation requirements [1, 16]. Teamwork in academic settings can positively impact students' creativity, students' understanding of complex concepts [22, 23], students' engagement, leadership, problem-solving skills, critical thinking, retention [17], and motivation in a positive way [24-27]. Developing teamwork skills for engineering students is essential because it can directly influence students' performance in their future careers, where employers increasingly value teamwork skills and the ability to perform well in teams [4, 28, 29]. As a result, team-based projects have become a feature of engineering curricula, particularly at the foundation level, specifically the first year [30, 31].

2.2 Teamwork Assessment

Given the importance of developing teamwork skills, the question of how to assess those skills becomes central. Peer assessment has been widely recognized for enhancing student learning by boosting motivation [32, 33], engagement [34], and critical thinking [35]. Serving as an assessor requires higher-order skills, such as reasoning and argumentation [36], which deepens understanding and encourages reflection [37]. It also promotes autonomy, professional growth, and confidence [38], while promoting interaction with peers and instructors [39]. In engineering education, since teamwork often occurs outside formal class settings where instructors cannot directly observe individual contributions, peer assessment becomes a valuable assessment tool when thoughtfully designed and implemented. Therefore, the overall effectiveness largely relies on the quality of the assessment instrument itself.

Although designing an effective assessment of any learning outcome is critical to learning, evaluating teamwork skills, as a process-based outcome [40], can be challenging when instructors have limited experience and resources. ABET suggests assessing teamwork skills through group projects that evaluate deliverables and teamwork processes, peer assessments of

contributions and communication within the team, self-reflection, and faculty observations during class activities [1]. However, instructors more often rely on indirect assessments, such as reflections on team experiences, to engage students in recognizing teamwork skills (e.g., [41]).

Some of the developed quantitative instruments have been critiqued for lacking a robust theoretical framework or desirable psychometric properties [42]. CATME is a widely used tool for assessing teamwork that relies on both self and peer evaluations [43]. However, it may not be affordable for all institutions due to its commercialization.

Another instrument for evaluating team effectiveness is grounded in the model originally developed by Imbrie, et al. [10], which is also available and in use. The first version of the model was based on studies on effective teamwork in industries. It identified three key factors that make teams successful: Interdependency, Goal setting, and Potency [10, 44]. A nine-item peer evaluation tool was developed to assess effective teamwork in the areas of interdependency, goal setting, and potency. Data were gathered from peer evaluations, and confirmatory factor analysis (CFA) was performed to evaluate construct validity.

On the other hand, Trust is considered an important factor for team effectiveness by many organizational researchers [45-50]. Trust was subsequently added as the fourth construct to the instrument. The validity and reliability of the instrument to assess four constructs have been thoroughly analyzed and confirmed [51]. However, evolving generational dynamics and the presence of hybrid teams, which often engage in virtual meetings and face diverse situations, motivated us to conduct this study to provide construct validity evidence and confirm the instrument's reliability, using a different sample, once more.

3. Theoretical Framework

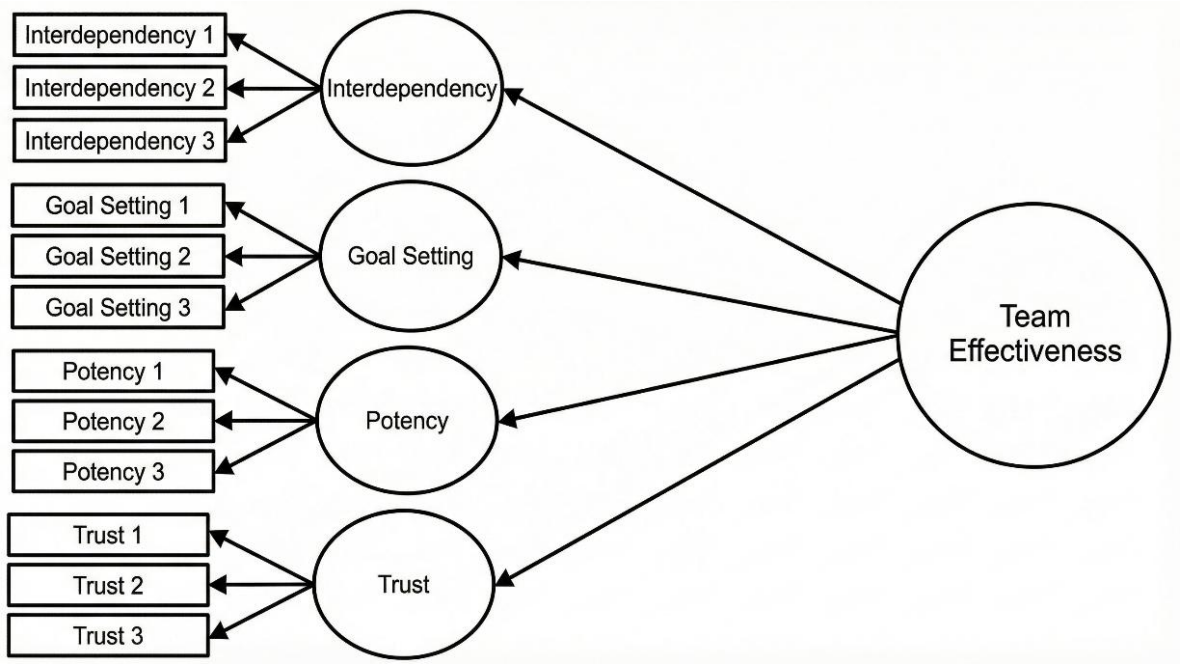
Given the importance of interdependency, goal setting, potency, and trust in evaluating student team effectiveness in engineering education [10, 11], this study employed the TDPES, a peer evaluation instrument developed by Wang and Imbrie [11], which is grounded in the four-factor team effectiveness model. Most importantly, the specific measurement of trust was considered an essential feature. Trust is a core ingredient for any successful team, yet other evaluation tools often overlook this element. Therefore, based on its cost-effectiveness, the rich dataset available, and its focus on this crucial element, it was determined to be the most suitable choice for the study. Each construct measured by the TDPES is defined in Table 1.

Table 1. Definitions of the Four Constructs Measured by the TDPES

Construct	Definition	References
Interdependency	The degree to which team members rely on one another to complete tasks and achieve goals.	[30, 52]
Goal setting	The process by which teams establish clear, specific, and shared objectives, serving as a critical transition activity for reflecting on past performance and planning future efforts.	[31, 46]
Potency	The team's collective belief in its ability to succeed and a shared confidence that goes beyond the sum of individual self-efficacy.	[52]

Trust	The willingness to be vulnerable to teammates based on positive expectations of their intentions and behaviors acting as the relational foundation that holds the team together under interdependence and risk.	[46, 53]

Figure 1. The Hypothesized Structure of the Four-Factor Model of the TPDES



4. Method

4.1 Research Context and Participants

The study was conducted within a required fall semester foundational engineering course for the first-year engineering program at a large, public, urban R1 university in the Midwestern United States in 2023 and 2024. A central feature of the course was a semester-long design project that required students to work in teams of three or four members. These teams were formed by the instructional staff at the beginning of the semester with the explicit goal of balancing member characteristics to ensure diversity in prior knowledge, experience, and demographics [54]. The course was taught in a flipped classroom format, integrating foundational engineering topics with hands-on project work. The total number of initial participants was 2,329. The final dataset for this study comprised 2,077 students, as shown in Table 1, which included a total of 667 teams after data cleaning and the removal of incomplete survey responses. Only incomplete responses were removed from the dataset using listwise deletion. Although demographic subgroups were not analyzed in this study, Table 2 presents the sample compositions to support readers' judgments about the transferability of the findings to similar first-year engineering populations.

Table 2. Demographic Profile of the 2,077 First-Year Engineering Students

Category	Subcategory	*N*	*%*
Gender	Female	465	22.4
	Male	1,608	77.4
	Unspecified	4	0.2
Race/ Ethnicity	American Indian or Alaska Native	3	0.1
	Asian	126	6.1
	Black	102	4.9
	Hispanic/Latino	101	4.9
	White	1,488	71.6
	Multi-racial	79	3.8
	Non-resident (International student)	162	7.8
	Unspecified	17	0.8
	Total	**2,077**	**100.0**

4.2 Data Collection and Instrument

Following the IRB approval, the data were collected using an end-of-semester survey administered during the Fall 2023 and Fall 2024 semesters. The survey instrument prompted students to perform a comprehensive peer evaluation of each teammate and a self-evaluation, using a combination of a scale, TDPES, and two open-ended questions. For this study to answer research questions, we only analyzed the TDPES data.

Team Development Peer Evaluation Scale (TDPES). The TDPES was selected over other alternatives for several key reasons. While other popular tools, such as CATME, can be expensive, this survey offers a low-cost alternative. It has also been adopted by several universities, which means a large and varied dataset is available for analysis, making it more likely that the findings are relevant beyond a single campus. The instrument's core is a peer evaluation tool built on a four-factor model of team effectiveness. It includes four factors, each with three items, to measure Interdependency, Goal Setting, Potency, and Trust. Students responded to each item using a five-point Likert scale (1 = Strongly Disagree, 2= Disagree, 3 = Neither Disagree nor Agree, 4 = Agree, 5 = Strongly Agree). As shown in Table 3, the instrument comprises four constructs, each with three items.

Table 3. 12 Items Aligned with the Four-Factor Model on the TDPES

Construct	Items	Descriptions
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 role
	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.

4.3 Data Analysis

The following analyses were conducted to re-evaluate the construct validity evidence of the instrument used for the study: confirmatory factor analysis (CFA) to assess construct validity and internal consistency analysis to assess reliability. Given that the data consists of ordered categorical variables (five-point Likert scale), statistical analyses were conducted using *R 4.5.3* with the *lavaan* package for the CFA and the *psych* package for reliability coefficients.

Utilizing the fit indices, we evaluated the fit of the CFA model by examining the consistency of multiple goodness-of-fit indices: the Chi-square (χ^2) statistic, the root mean square error of approximation (RMSEA), the comparative fit index (CFI), and the Tucker-Lewis index (TLI) [55]. Additionally, we computed Cronbach's α as a measure of internal consistency reliability to assess the interrelationships among items within each factor and the overall instrument.

5. Results

5.1 Construct Validity Evidence

The CFA model fit was evaluated by examining the consistency of multiple goodness-of-fit indices: the Chi-square (χ^2) test, the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), and the Tucker-Lewis Index (TLI). The model specified for the four-factor structure yielded the following fit indices: $\chi^2(50) = 651.93$, $p < .001$, CFI = .977, TLI = .970, RMSEA = .076, 90% CI (.071, .081). Although the Chi-square test was significant, which is common in large samples [56], the CFI (.977) and TLI (.970) exceeded the recommended threshold of .95 [57]. The RMSEA value of .076, 90% CI (.071, .081), fell within the acceptable range of .05 to .08 recommended by Browne and Cudeck [58], though it exceeded the more stringent .06 cutoff proposed by Hu and Bentler [53]. This is consistent with the known sensitivity of RMSEA to model complexity in large samples [59]. Taken together, the pattern of fit indices supported the adequacy of the four-factor model for this sample. All factor loadings were statistically significant ($p < .001$), confirming that each observed item loaded onto its designated latent factor as specified by the model [60]. Taken together, these results provide evidence of construct validity for the four-factor model within this sample of first-year engineering students, which contributes one strand of evidence to the broader construct validity

argument for the TDPES.

5.2 Reliability Evidence

Data from the full student sample (n = 2,077) were utilized for the reliability analysis. The internal consistency reliability for each of the four constructs and the overall TDPES was assessed by calculating Cronbach's α. The reliability coefficients were .924 for Interdependency, .914 for Goal Setting, .923 for Potency, and .863 for Trust. Cronbach's α for the overall TDPES was .962. All coefficients were in an acceptable to excellent range, indicating sound internal consistency reliability evidence for the TDPES [60, 61].

6. Discussion

The purpose of this study was to provide evidence of construct validity and reliability of the four-factor peer evaluation instrument for a contemporary cohort of first-year engineering students. The results from the CFA provided evidence supporting the hypothesized structure of the instrument originally developed by Wang and Imbrie [11]. The model fit indices, χ^2 = 651.93, CFI = .977, TLI = .970, RMSEA = .076, met the established thresholds for acceptable model fit, with CFI and TLI exceeding the recommended .95 threshold and RMSEA falling within the acceptable-but-not-excellent range below .08 [55]. This confirms that the four latent factors, Interdependency, Goal Setting, Potency, and Trust, are distinct yet related constructs that accurately reflect the underlying data.

This finding aligns with previous validations of the instrument and suggests that the model remains robust despite evolving generational dynamics and the increasing prevalence of hybrid team interactions. Furthermore, the internal consistency reliability for all four constructs was high, with Cronbach's α coefficients ranging from .863 to .924. These coefficients fall within the acceptable to excellent range, providing evidence of internal consistency for the instrument [61, 62]. Notably, the "Trust" construct demonstrated strong reliability (α = .863) and significant factor loadings. This provides evidence of validity to include Trust as a critical component of team effectiveness, distinguishing this instrument from others that may neglect the relational aspects of teamwork emphasized by organizational researchers [46, 53].

6.1 Limitations and Future Research

Several limitations should be considered when interpreting these findings. First, the present study contributes evidence of construct validity and reliability, which is one strand of a broader construct validity argument. A complete validity case for the TDPES requires additional types of evidence, including content evidence, response process evidence, relations with other variables such as team performance and individual academic outcomes, measurement invariance across gender, race, and team composition, and evidence of consequences of score use. Future research should pursue these strands to build a more comprehensive construct validity argument for the TDPES.

Second, the data were collected at a single institution from first-year engineering students, which limits the generalizability of the findings to other populations, such as upper-level students,

capstone teams, or students in specific engineering disciplines and institutional contexts. Third, peer evaluations reflect students' perceptions of their teammates rather than direct behavioral observation, and may be influenced by friendship, social desirability, or leniency biases. Fourth, the RMSEA value of .076 falls within the acceptable range proposed by Browne et al. [54] but exceeds the more stringent .06 threshold proposed by Hu and Bentler [53], indicating acceptable rather than excellent approximate fit. Together, these limitations point to clear directions for future work, including replication with upper-level and capstone student samples, invariance testing across demographic subgroups, and linking TDPES scores to behavioral and performance outcomes.

6.2 Conclusion

This study aimed to re-evaluate the construct validity evidence of the TDPES, a peer evaluation instrument designed to measure team effectiveness in engineering education. The results from the CFA provided robust evidence supporting the four-factor structure of Interdependency, Goal Setting, Potency, and Trust. The fit indices fell well within acceptable ranges, and the reliability coefficients indicated strong internal consistency for all constructs. These findings suggest that including "Trust" as a distinct factor is statistically justified and essential for a holistic view of team dynamics. By confirming the stability of these factors, educators can confidently use the results to identify specific areas of team dysfunction and implement targeted interventions to support team development.

The study has implications for both research and practice in engineering education. For practitioners, this instrument offers a validated, cost-effective alternative to commercial platforms for assessing student teams. For researchers, it provides a reliable metric to track how specific interventions, such as those targeting trust or goal setting, influence team performance over time. Future work will focus on using the TDPES to develop a targeted skill building system, allowing instructors to provide specific interventions based on team scores. Ultimately, this tool empowers educators to move beyond simple assessment and toward actively developing students' professional competencies.

Generative AI Disclosure

During the preparation of this work, the authors used Grammarly AI to check for grammatical errors and improve the writing flow and consistency. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the accuracy and integrity of the final publication.

Reference

[1] ABET. "Criteria for Accrediting Applied and Natural Science Programs, 2025-2026." https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-applied-and-natural-science-programs-2025-2026/ (accessed Aug. 20, 2025.

[2] C. Winberg *et al.*, "Developing employability in engineering education: a systematic review of the literature," *European Journal of Engineering Education,* vol. 45, no. 2, pp. 165–180, 2020.

[3] E. Britton, N. Simper, A. Leger, and J. Stephenson, "Assessing teamwork in undergraduate education: a measurement tool to evaluate individual teamwork skills," *Assessment & Evaluation in Higher Education,* vol. 42, no. 3, pp. 378–397, 2015, doi: 10.1080/02602938.2015.1116497.

[4] J. Paoletti, T. M. Bisbey, D. L. Reyes, M. A. Wettergreen, and E. Salas, "A checklist to diagnose teamwork in engineering education," *International Journal of Engineering Education,* vol. 36, no. 1, pp. 365–377, 2020.

[5] A. R. Rajabzadeh, J. Long, G. Saini, and M. Zeadin, "Engineering Student Experiences of Group Work," *Education Sciences,* vol. 12, no. 5, p. 288, 2022, doi: 10.3390/educsci12050288.

[6] M. L. Loughry, M. W. Ohland, and D. DeWayne Moore, "Development of a Theory-Based Assessment of Team Member Effectiveness," (in English), *Educational and Psychological Measurement,* vol. 67, no. 3, pp. 505–524, Jun 2007, doi: 10.1177/0013164406292085.

[7] T. O'Neill *et al.*, "Introducing a scalable peer feedback system for learning teams," *Assessment & Evaluation in Higher Education,* Article vol. 44, no. 6, pp. 848–862, 2019, doi: 10.1080/02602938.2018.1526256.

[8] A. Warren, "TEAMMATES peer feedback system – review," ed, 2015.

[9] Iconic Learning Systems LLC. "Team+| Teams that work." https://www.teamplus.education/about-us/ (accessed Jul. 18, 2023.

[10] P. Imbrie, S. Maller, and J. Immekus, "Assessing team effectiveness," in *2005 Annual Conference*, 2005, pp. 10.229. 1–10.229. 7.

[11] J. Wang and P. K. Imbrie, "Assessing Team Effectiveness: Comparing Peer Evaluations To A Team Effectiveness Instrument," June 2009, Austin, Texas: ASEE Conferences, pp. 14.249.1–14.249.9, doi: 10.18260/1-2--5770. [Online]. Available: http://peer.asee.org/5770

[12] A. G. Erickson, P. Noonan, K. S. Carter, L. McGurn, and E. Purifoy, "The team functioning scale: Evaluating and improving effectiveness of school teams," *International Journal of Educational Research,* vol. 69, pp. 1–11, 2015.

[13] J. R. Katzenbach and D. K. Smith, "The rules for managing cross-functional reengineering teams," *Planning review,* vol. 21, no. 2, pp. 12–13, 1993.

[14] E. Salas, T. L. Dickinson, S. A. Converse, and S. I. Tannenbaum, "Toward an understanding of team performance and training," in *Teams: Their training and performance*, R. W. S. a. E. Salas Ed. Norwood, NJ: Ablex, 1992, pp. 3–29.

[15] E. Salas, D. L. Reyes, and A. L. Woods, "The Assessment of Team Performance: Observations and Needs," in *Innovative Assessment of Collaboration*, (Methodology of Educational Measurement and Assessment, 2017, ch. Chapter 2, pp. 21–36.

[16] J. E. Mathieu, P. T. Gallagher, M. A. Domingo, and E. A. Klock, "Embracing Complexity: Reviewing the Past Decade of Team Effectiveness Research," *Annual Review of Organizational Psychology and Organizational Behavior,* vol. 6, no. 1, pp. 17–46, 2019, doi: 10.1146/annurev-orgpsych-012218-015106.

[17] G. R. Goldsmith *et al.*, "Overcoming the Barriers to Teaching Teamwork to Undergraduates in STEM," *CBE Life Sci Educ,* vol. 23, no. 2, p. es2, Jun 2024, doi: 10.1187/cbe.23-07-0128.

[18] A. A. Von Davier, M. Zhu, and P. C. Kyllonen, *Innovative assessment of collaboration*. Springer, 2017.

[19] J. R. Hackman and N. Katz, "Group behavior and performance," *Handbook of social psychology,* vol. 2, no. 1208-1251, p. 2, 2010.
[20] R. Grossman, S. Friedman, and S. Kalra, "The Wiley Blackwell Handbook of the Psychology of Team Working and Collaborative Processes," 2017.
[21] J. E. Mathieu, J. R. Hollenbeck, D. van Knippenberg, and D. R. Ilgen, "A century of work teams in the Journal of Applied Psychology," *Journal of Applied Psychology,* vol. 102, no. 3, pp. 452–467, Mar 2017, doi: 10.1037/apl0000128.
[22] R. Reiter-Palmon and J. J. Illies, "Leadership and creativity: Understanding leadership from a creative problem-solving perspective," *The Leadership Quarterly,* vol. 15, no. 1, pp. 55–77, 2004/02// 2004, doi: 10.1016/j.leaqua.2003.12.005.
[23] R. E. Slavin, "Cooperative Learning and Academic Achievement: Why Does Groupwork Work?. [Aprendizaje cooperativo y rendimiento académico: ¿por qué funciona el trabajo en grupo?]," (in en), *analesps,* vol. 30, no. 3, pp. 785–791, 2014/10/01/ 2014, doi: 10.6018/analesps.30.3.201201.
[24] 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," *Organizational Research Methods,* vol. 23, no. 3, pp. 399–431, 2020. [Online]. Available: https://journals-sagepub-com.uc.idm.oclc.org/doi/full/10.1177/1094428119840801.
[25] J. R. Hackman, "From causes to conditions in group research," *Journal of Organizational Behavior,* vol. 33, no. 3, pp. 428–444, 2012, doi: 10.1002/job.1774.
[26] Y. Chang and P. Brickman, "When Group Work Doesn't Work: Insights from Students," (in en), *CBE Life Sci Educ,* vol. 17, no. 3, p. ar42, Sep 2018, doi: 10.1187/cbe.17-09-0199.
[27] T. Rapp, T. Maynard, M. Domingo, and E. Klock, "Team Emergent States: What Has Emerged in The Literature Over 20 Years," *Small Group Research,* vol. 52, no. 1, pp. 68–102, 2020, doi: 10.1177/1046496420956715.
[28] E. Salas *et al.*, "Does team training improve team performance? A meta-analysis," *Human Factors,* vol. 50, no. 6, pp. 903–33, Dec 2008, doi: 10.1518/001872008X375009.
[29] 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," (in en), *Academy of Management Learning & Education,* vol. 16, no. 4, pp. 562–578, 2017/12// 2017, doi: 10.5465/amle.2016.0163.
[30] 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, 2013, doi: 10.1002/jee.20023.
[31] H. G. Murzi, T. M. Chowdhury, J. Karlovšek, and B. R. Ulloa, "Working in large teams: Measuring the impact of a teamwork model to facilitate teamwork development in engineering students working in a real project," *International Journal of Engineering Education,* vol. 36, no. 1, pp. 274–295, 2020.
[32] S. J. Deeley and C. Bovill, "Staff student partnership in assessment: Enhancing assessment literacy through democratic practices," *Assessment & Evaluation in Higher Education,* vol. 42, no. 3, pp. 463–477, 2017.
[33] J. H. Stonewall, "Development and evaluation of training to increase student perceptions of fairness in peer assessment," Iowa State University, 2022.

[34] S. Bloxham and A. West, "Understanding the rules of the game: marking peer assessment as a medium for developing students' conceptions of assessment," *Assessment & Evaluation in Higher Education,* vol. 29, no. 6, pp. 721–733, 2004.
[35] T. Harland, N. Wald, and H. Randhawa, "Student peer review: Enhancing formative feedback with a rebuttal," *Assessment & Evaluation in Higher Education,* vol. 42, no. 5, pp. 801–811, 2017.
[36] K. J. Topping, "Peers as a source of formative and summative assessment," in *SAGE handbook of research on classroom assessment*: Sage Publications, Inc., 2013, pp. 394–412.
[37] J. D. Snowball and M. Mostert, "Dancing with the devil: Formative peer assessment and academic performance," *Higher Education Research & Development,* vol. 32, no. 4, pp. 646–659, 2013.
[38] K. Topping, "Peer assessment between students in colleges and universities," (in English), *Review of educational Research,* vol. 68, no. 3, pp. 249–276, Fal 1998, doi: 10.2307/1170598.
[39] M. L. Wen and C.-C. Tsai, "University students' perceptions of and attitudes toward (online) peer assessment," *Higher Education,* vol. 51, no. 1, pp. 27–44, 2006. [Online]. Available: https://link-springer-com.uc.idm.oclc.org/content/pdf/10.1007/s10734-004-6375-8.pdf.
[40] E. R. Kemery and L. T. Stickney, "A Multifaceted Approach to Teamwork Assessment in an Undergraduate Business Program," *Journal of Management Education,* vol. 38, no. 3, pp. 462–479, 2013, doi: 10.1177/1052562913504762.
[41] A. K. Barker, "Developing teamwork skills in a course on ore deposits," *Mineralogia,* vol. 52, no. 1, pp. 43–48, 2022, doi: 10.2478/mipo-2021-0005.
[42] B. Vaughan, J. Yoxall, and S. Grace, "Peer assessment of teamwork in group projects: Evaluation of a rubric," *Issues in Educational Research,* vol. 29, no. 3, pp. 961–978, 2019.
[43] M. W. Ohland *et al.*, "The Comprehensive Assessment of Team Member Effectiveness: Development of a Behaviorally Anchored Rating Scale for Self- and Peer Evaluation," (in English), *Academy of Management Learning & Education,* vol. 11, no. 4, pp. 609–630, Dec 2012, doi: 10.5465/amle.2010.0177.
[44] T. J. Moore, H. Diefes-Dux, and P. K. Imbrie, "Assessment of Team Effectiveness During Complex Mathematical Modeling Tasks," in *Proceedings. Frontiers in Education. 36th Annual Conference*, 27–31 Oct. 2006, pp. 1–6, doi: 10.1109/FIE.2006.322481. [Online]. Available: https://ieeexplore.ieee.org/stampPDF/getPDF.jsp?tp=&arnumber=4117015&ref=
[45] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An integrative model of organizational trust," *Academy of management review,* vol. 20, no. 3, pp. 709–734, 1995.
[46] M. A. Marks, J. E. Mathieu, and S. J. Zaccaro, "A temporally based framework and taxonomy of team processes," (in English), *Academy of management review,* vol. 26, no. 3, pp. 356–376, 2001, doi: 10.5465/Amr.2001.4845785.
[47] F. Ranjbar, J. Agarwal, E. Vahidi, J. Wang, and P. Imbrie, "WIP: A Preliminary Investigation of Students as Peer Evaluators of Their Team Members," in *2024 IEEE Frontiers in Education Conference (FIE)*, 2024: IEEE, pp. 1–5.

[48] F. Ranjbar, J. Wang, S. A. Khanamani, J. Agarwal, and P. Imbrie, "WIP: Negative and Positive Behaviors Exhibited by Students in the Context of Engineering Project Team," in *2025 IEEE Frontiers in Education Conference (FIE)*, 2025: IEEE, pp. 1–5.

[49] J. Wang, F. Ranjbar, and P. Imbrie, "WIP: Students in Dysfunctional Team: First Year Engineering Design Course Perspective," in *2025 IEEE Frontiers in Education Conference (FIE)*, 2025: IEEE, pp. 1–5.

[50] J. Wang, F. Ranjbar, and P. Imbrie, "WIP: Teaming Performance as Measured by Peer Evaluations: A Comparison Between Hyflex Team and Face to Face Team," in *2025 IEEE Frontiers in Education Conference (FIE)*, 2025: IEEE, pp. 1–5.

[51] J. Wang, "Assessment of engineering student team effectiveness," Doctoral dissertation, Engineering Education, Purdue University, 2014.

[52] S. M. Gully, K. A. Incalcaterra, A. Joshi, and J. M. Beaubien, "A meta-analysis of team-efficacy, potency, and performance: interdependence and level of analysis as moderators of observed relationships," *Journal of Applied Psychology,* vol. 87, no. 5, pp. 819–32, Oct 2002, doi: 10.1037/0021-9010.87.5.819.

[53] B. A. De Jong, K. T. Dirks, and N. Gillespie, "Trust and team performance: A meta-analysis of main effects, moderators, and covariates," *Journal of Applied Psychology,* vol. 101, no. 8, pp. 1134–50, Aug 2016, doi: 10.1037/apl0000110.

[54] 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)*, 8–11 Oct. 2022 2022, pp. 1–6, doi: 10.1109/FIE56618.2022.9962741. [Online]. Available: https://ieeexplore.ieee.org/stampPDF/getPDF.jsp?tp=&arnumber=9962741&ref=

[55] T. A. Brown, *Confirmatory factor analysis for applied research*. Guilford publications, 2015.

[56] R. B. Kline, *Principles and practice of structural equation modeling*. Guilford publications, 2023.

[57] L. Hu and P. M. Bentler, "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives," (in English), *Structural equation modeling: a multidisciplinary journal,* vol. 6, no. 1, pp. 1–55, 1999, doi: 10.1080/10705519909540118.

[58] M. W. Browne and R. Cudeck, "Alternative ways of assessing model fit," in *Testing Structural Equation Models*, K. A. B. a. J. S. Long Ed. Newbury Park, CA: Sage, 1993, pp. 136–162.

[59] F. Chen, P. J. Curran, K. A. Bollen, J. Kirby, and P. Paxton, "An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models," *Sociological methods & research,* vol. 36, no. 4, pp. 462–494, 2008. [Online]. Available: https://journals-sagepub-com.uc.idm.oclc.org/doi/pdf/10.1177/0049124108314720?download=true.

[60] J. C. Nunnally and I. H. Bernstein, *Psychometric Theory*, 3rd ed. New York, NY: McGraw-Hill, 1994.

[61] A. Field, *Discovering statistics using IBM SPSS statistics*. sage, 2013.

[62] M. J. Stevens and M. A. Campion, "Staffing Work Teams: Development and Validation of a Selection Test for Teamwork Settings," *Journal of Management,* vol. 25, no. 2, pp. 207–228, 1999, doi: 10.1177/014920639902500205.