Peer Evaluation of Teamwork Activities in Undergraduate Engineering Design

Jiachun Huang[a], Scott Wordley[a], Ashlee Pearson[a]
Department of Mechanical and Aerospace Engineering, Monash University[a]
Corresponding Author Email: jiachun.huang@monash.edu

SESSION C3: Integration of teaching and research in the engineering training process

CONTEXT This paper reports that a formal peer evaluation system combined with teamwork development activities may be an option to facilitate professional skill development in undergraduate students. This paper outlines the implementation of an online, peer evaluation system designed to facilitate feedback, mark moderation and student reflection on teamwork. This system is also helpful in discouraging poor team conduct by making students accountable to their teammates for the quality and quantity of work that they contribute and their communication in team-based assessments. Self-evaluation is also used to encourage reflective practice and highlight any differences in perceived contributions.

PURPOSE The objective of this research is to assess student attitudes towards the use of an anonymous, online peer evaluation tool (CATME) and teamwork and professional skill development in a second year engineering design unit, where team-based assessments make up a significant portion of the overall grade.

APPROACH Multimedia resources and team meeting activities were developed to introduce CATME to students and to support their use of the tool throughout the semester. An anonymous online survey was administered at the end of the semester to assess student attitudes towards the use of a formal peer evaluation system and formal team development activities. Students were also asked to rate their level of satisfaction with their team and how the activities within the unit contributed to the development of teamwork and interpersonal skills. An analysis of open ended written comments submitted as part of the survey was used to provide additional insight into student attitudes and perceptions.

RESULTS Findings from the online survey show that on average, respondents agreed that they should evaluate the contribution of their peers, that they felt comfortable knowing their marks would be moderated based on peer evaluation and that a formal peer evaluation system was a fair way to moderate team marks. By the end of the unit, students agreed that their interpersonal communication, ability to resolve conflicts, management skills and accountability had improved. There was also agreement that the formal team development activities were helpful in facilitating teamwork in aspects such as communication and project planning. On average, students indicated that they would rather work in teamwork activities and team projects in which there is a formal peer evaluation system. Open ended written comments indicated that students appreciated having a formal system which enforces individual accountability, and through which they could report on the contribution of individual team members, knowing that they were supported by a fair process of mark moderation.

CONCLUSIONS An online peer evaluation system was successfully implemented in a second year engineering design unit; its use was accompanied by various multimedia resources and formal team development activities. Survey results suggest that students responded favourably to this system, and as a result it will be deployed in more engineering design units in the coming years.

KEYWORDS CATME; peer evaluation; formative assessment; reflective practice

Introduction

Team-based assessments can help students to develop the knowledge, skills and abilities (KSAs) that increase students' graduate employability. Stevens and Campion (1994) outline a set of KSA requirements for teamwork that include: conflict resolution, collaborative problem solving, communication, goal setting, planning and coordination. These professional teamwork and communication skills are required in graduate engineers (Engineers Australia 2011) and are highly regarded by graduate employers (May and Strong, 2011). However, students working in teams can encounter various teamwork issues, a common cause of which is social loafing, a phenomenon where individuals exert less effort when working in a group or team (Simms and Nichols, 2014). Teamwork issues can manifest as breakdowns in communication, poor decision making and task and relational conflict (Borrego et al. 2013). When left unresolved and without instructor support, these issues can negatively affect students' satisfaction with teamwork. Oakley et al. (2007) surveyed over six thousand students in Engineering and found that student satisfaction in teams is significantly affected by the presence of 'slackers' and the level of instructor guidance on teamwork. This finding echoes that of Bolton (1999), who found higher satisfaction rates when students were provided with additional team training and instructor support.

Peer evaluation (PE) has been suggested as an effective option for mitigating social loafing and reducing the occurrence of conflict (Albanese and Van Fleet, 1985; Chen and Lou, 2004; Pfaff and Huddleston, 2003). PE enables instructors to manage and moderate teamwork by allowing students to evaluate themselves and their peers in aspects of teamwork such as technical ability and team dynamics. PE can be used to encourage effective teamwork by rewarding effective team members and discouraging poor teamwork by adjusting individual grades to reflect teamwork contribution (Ohland et al., 2012; Brutus and Donia, 2010; Kaufman et al., 2000; Brown, 1995). There is also a growing body of research that suggests PE as an effective means of developing teamwork KSAs (Brutus and Donia, 2010; Willey and Gardner, 2007; Loughry et al., 2014). PE may be used to improve students' teamwork satisfaction in two ways: 1) rewarding effective team members and associating poor teamwork with negative consequences on student grades (mark moderation) and 2) facilitating the development of teamwork KSAs through feedback and reflection.

There are various online peer evaluation platforms; some examples include: CATME (Ohland et al., 2012); SPARKPLUS (Willey and Gardner, 2010) and WebPA (Loddington et al., 2009). CATME (Comprehensive Assessment of Team Member Effectiveness) is a platform that can be used for both mark moderation and for facilitating the development of teamwork KSAs through feedback and reflection. CATME Peer Evaluation is used in this study owing to its Behaviourally Anchored Rating Scale (BARS) evaluation rubric (Ohland et al., 2012) and the simple interface. CATME assesses team member effectiveness in five main categories: 1) Contributing to the team's work; 2) Interacting with teammates; 3) Keeping the team on track; 4) Expecting quality; 5) Having relevant knowledge skills and abilities. The CATME peer evaluation survey requires students to rate themselves and their team members on these aspects of teamwork according the BARS evaluation rubric. CATME peer evaluation facilitates mark moderation of team projects using an adjustment factor, which is computed as:

$$\text{Adj Factor} = \frac{average\ rating\ of\ the\ student}{overall\ average\ rating\ for\ the\ team}$$

Mark moderation involves multiplying the adjustment factor with the mark received by the team for the assessment to generate an individual student mark. High-performing team members are rewarded while poor-performing or non-contributing members may be scaled down. This process helps instructors moderate teamwork so that students receive an individual mark that reflects their contributions to the team.

CATME peer evaluation was implemented in a second year Engineering Design unit at a large university in Australia. Students also participated in formal team development activities to supplement the peer evaluation surveys. The aims of this study are to investigate student attitudes towards using a formal peer evaluation system and to examine the effect of peer evaluation and team development activities on students' teamwork satisfaction and their perception of learning outcomes.

Theoretical Perspective

The approach to facilitating professional skills development using peer evaluation and team development activities is largely informed by Schön's (1983) theory of reflective practice and Sadler's (1989) theory of formative assessment. In Schön's work, reflection is described as a means by which a practitioner, through experience, builds upon their knowledge base to reach a better understanding of their actions and behaviours and inform future actions and decision-making. Loughran (2002) adds to Schön's theory by asserting that experience alone does not lead to learning and that more essentially it is the reflection on experience that can enhance learning. That is, teamwork alone may not help students develop professional skills; undergraduate students in particular need opportunities to reflect on teamwork experiences, emotions and outcomes as a way to draw insight and understanding of teamwork to inform future actions.

Formative assessment, where feedback is the main mechanism for learning, can facilitate reflective thinking. In the context of teamwork, Sadler's prescription of the three conditions for feedback are: 1) an understanding of what constitutes effective teamwork; 2) an ability to evaluate team member effectiveness against a standard and 3) the tools and knowledge to improve performance. The first and second points outlined by Sadler are satisfied by the BARS in CATME peer evaluation and the third condition is met with the team development activities embedded in the unit design.

Peer evaluation is a practice through which students have the opportunity to regularly reflect on teamwork experiences and interpersonal interactions. When used formatively, it can enable students to gain knowledge of the standards (building their knowledge base) for good teamwork and develop students' ability to evaluate the performance of their peers and themselves. Used summatively (mark moderation), peer evaluation can elicit from students a sense of individual accountability for teamwork (Kaufman et al., 2000; Brown, 1995). Finally, the feedback process is completed when students engage in self and team reflection which is facilitated by team development activities. Schön and Sadler's theories inform the pedagogical practices described in this study and the existing literature suggests that peer evaluation can improve student satisfaction by discouraging poor teamwork behaviour while team development activities facilitate effective teamwork.

Context

In semester one of 2015, a trial was conducted with a peer evaluation tool in a second year Mechanical, Aerospace and Mechatronics Engineering design unit with 285 student enrolments. The major continuous assessment in this unit involved students working in teams of four to five to design, build and test an autonomous device to navigate a track and deliver a payload (Smith, 2008). Heterogeneous teams were formed by the instructor with the intent to create a good mix of male/female students, cultural diversity, academic performance, international/domestic status and hands-on building experience (Felder and Brent, 2001). Students remained in their assigned teams for the duration of the major team project and semester (12 weeks). Teams made a preliminary submission for their design in week five, competed with other teams in a campus competition in week nine and submitted a final team report in week twelve. Students completed two rounds of CATME peer evaluation surveys to evaluate teamwork - once after the preliminary submission and once after the final report submission which included consideration of the campus competition. The instructor requested

that all students provide at least one sentence for each team member (including themselves) to justify their ratings in the 'Comments to the Instructor' text field at the end of the survey. Students were warned that individual marks (up to 10%) could be deducted for failure to provide meaningful written comments on aspects of teamwork. Students were informed that they would receive individual marks for the two major project submissions based on the adjustment factor generated from the peer evaluation system, subject to moderation and manual changes made by the instructor. The instructor performed moderation based on the rating patterns within the team and written comments from each student about their team. Students' individual marks could be scaled up by a maximum of 10% of their team's mark, but there was no limit on the downside scaling. Students could not receive more than full marks for any project submission.

To support effective teamwork, scheduled class time was allocated for students to complete team development activities at four scheduled intervals throughout the semester. Students completed an icebreaker activity in week one, a project planning meeting in week three, a teamwork reflection meeting in week five and a peer evaluation feedback meeting in week six. The instructor provided students with agenda and meeting minutes templates to guide students in practicing effective management and communication. Teams were also provided with a short video (four minutes) introducing them to the CATME peer evaluation system. Students were required to practice evaluating a hypothetical team before evaluating their real team members in the first round. The system provided feedback to students on the accuracy of their ratings of the hypothetical team members which helped students to understand the BARS.

Method

At the end of the semester, students were asked to complete an online, anonymous survey to capture their attitudes toward the peer evaluation process and team development activities. Students responded to twenty-five questions on various aspects of teamwork in the unit including: mark moderation; peer evaluation; teamwork satisfaction; professional skills development and team development activities. Responses were recorded on a seven-point Likert scale between strongly disagree (-3) and strongly agree (+3) for all but one question which measured whether students felt that the mark moderations were too small (-3) or too large (+3). As part of the survey, students were also asked to provide written comments on their experience with using peer evaluations in the unit. The survey data was analysed with descriptive statistics (proportions, means and variances) and themes were uncovered in the written comments with simple qualitative coding. Ethics approval for this research was sought and granted from the university's human research ethics committee (project number: CF15/2901 - 2015001197).

Results & Discussion

The survey received a response rate of 24% (67 students). Demographic data such as gender, ethnicity and age were not recorded. This was a limitation in this study as there was little indication of whether the respondents were proportionately representative of the cohort. A histogram of the responses for each question is represented by shaded boxes where the darkness of shade indicates the number of counts for each position on the Likert scale (Figure 1). The distribution of responses for each question was displayed in this manner to simultaneously illustrate the skewness in the responses, the mean response, and to provide a good indication of the variance in the Likert scale ratings. Ratings adjacent to neutral (-1 and +1) are considered as relatively neutral, ratings between 1 and 2 are considered to be moderate agreement, and ratings of +2 and +3 are considered as strong agreement.

Variance in Responses

Students tended to be in agreement with the statements provided, with their responses exhibiting a negative skew (Figure 1). This could indicate that students truly felt positively about

teamwork, peer evaluation and team development activities, though it is also possible that the responses, like many others measured on Likert scales, tend to be negatively skewed (Peterson and Wilson, 1992). The small variance in questions 1 - *severity of mark moderation* (variance = 1.5), 21 - *helped my team get to know each other* (variance = 1.3), and 22 - *helped my team to establish lines of communication* (variance = 0.9) indicates that there is a strong consensus in the cohort that the mark moderations were neither too small or too large and that the team development activities facilitated effective communication early in the semester.
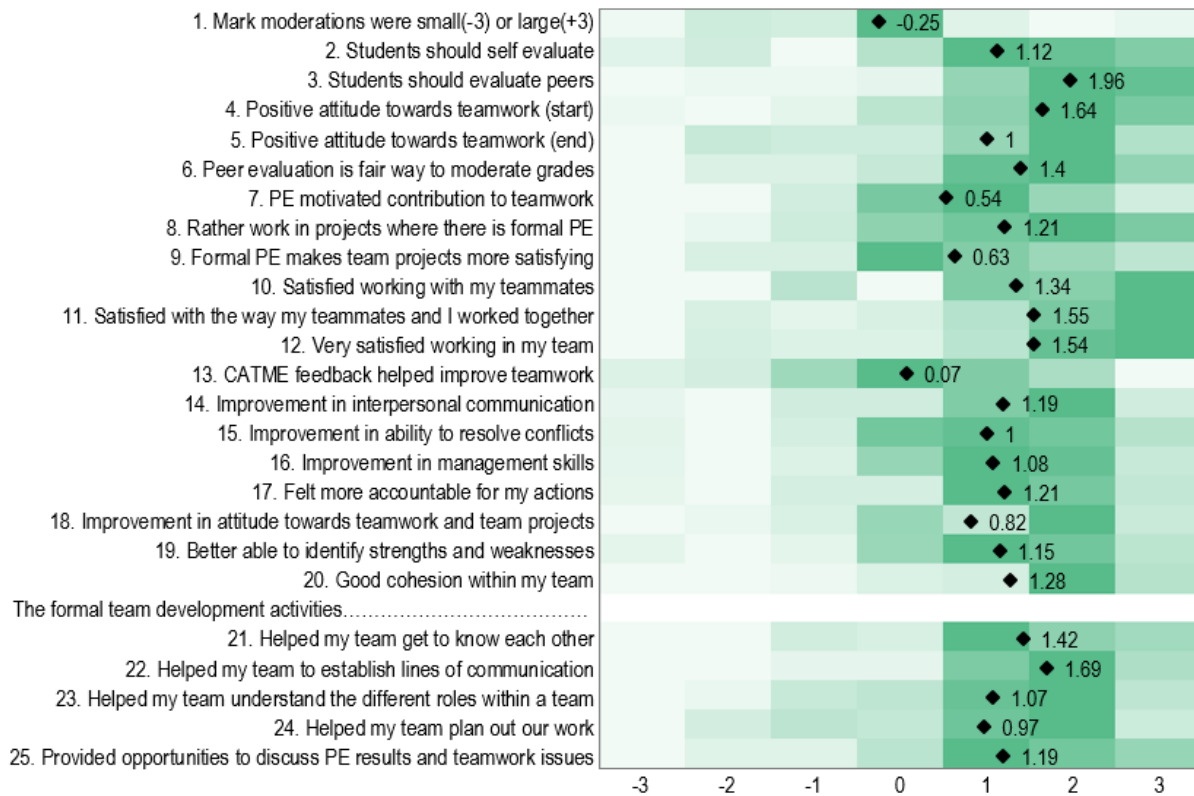

Figure 1: The average response (diamond) on a seven point Likert scale from strongly disagree (-3) to strongly agree (+3) unless otherwise specified. The data labels for the average is displayed numerically next to each diamond. The shaded region represents a histogram of the number of responses for each question; darker regions indicate a higher count of responses.

The variance was largest in question 2 - *students should self-evaluate* (variance = 2.7); the cohort was divided on the idea that students should self-evaluate their contributions to teamwork. There was greater consensus that *students should evaluate their peers on teamwork contribution* (variance = 1.1). This likely indicates that a considerable portion of the class failed to appreciate the importance of self-appraisal in comparison with peer-appraisal. In a cohort of undergraduate engineering students, this is not surprising; it's likely that these second-year students are yet to develop the reflective mindset that would allow them to appreciate the significance of self-evaluation. Repeated practice with self and peer evaluation may help students to develop a reflective mindset over time. The next largest variances were seen in questions 10 – *I was satisfied working with my teammates* (variance = 2.7) and 18 - *I felt that my attitude towards teamwork and team-based projects had improved* (variance = 2.7). Both variances indicate that the cohort was widely mixed with good and poor teamwork experiences throughout the semester. Again, this comes as no surprise; for many students this unit was the first in which they engaged significantly in teamwork. The combination of heterogeneous instructor-formed teams and an above average workload meant that each student's experience of teamwork depended heavily on how well they got along with their teammates. Inevitably, a small number of teams encountered conflict and discontent, leading to a relatively large variance in teamwork satisfaction and shift in teamwork attitudes.

Professional Skill Development

The survey data indicates that students on average felt they had improved in several professional skill categories, namely interpersonal communication, ability to resolve conflicts, management skills, accountability and ability to identify strengths and weaknesses (Q14-17, 19). A large proportion of the cohort responded with the second highest (+2) and highest (+3) agreement rating to these questions (Table 1).

Table 1: Percentage of the class in strong agreement for questions 14-17 and 19 (responses in the second-highest or highest agreement rating).

Question: At the end of the current unit...	Percentage
...I felt that my interpersonal communication had improved.	51%
...I felt that my ability to resolve conflicts within the team had improved.	38%
...I felt that my management skills had improved.	41%
...I felt more accountable for my actions.	45%
...I was better able to identify my strengths and weaknesses.	42%

Just over half the class indicated strong agreement that their interpersonal skills had improved; smaller though nonetheless considerable portions of the class indicated strong agreement that their management skills, ability to resolve conflicts and ability to identify own strengths and weaknesses had improved. It is likely that PE and formal team development activities played a role in these improvements; this could be confirmed by conducting a controlled study.

Accountability

Nearly half the class indicated strong agreement that at the end of the unit they felt more accountable for their actions (Table 1). Individual student marks were generated by adjusting the team's mark for project submissions based on students' peer evaluation adjustment factors. Students were informed about this practice at the beginning of the semester, which helped to ensure that students were fully aware that the marks they receive would be affected by their team contributions and team member effectiveness. It is likely that this practice of mark moderation played a significant role in establishing individual accountability amongst team members for their actions and behaviours in the team. In the first round of peer evaluation, adjustment factors were moderated by the instructor so that students received slightly higher adjustment factors than the raw values generated by CATME. This practice served as an initial warning to poor performing or non-contributing team members to improve their performance. In the second round of peer evaluation, the instructor was more critical with moderation of individual marks; one student even received an adjustment factor of zero (this student failed to engage with the team and project), meaning that they received an individual mark of zero for the final project submission.

Written Comments

Best aspects of the peer evaluation system

Analysis of the written comments regarding the best aspects of peer evaluation revealed that *fairness/accountability* and *feedback* were the prevailing themes. Students appreciated that there was a formal moderation process in place to reflect individual contributions to teamwork and that this process held team members accountable (22 comments). A notable comment that reflects this sentiment was: *"If you have to do more of the work than team members who are lazy/don't take any initiative, you get rewarded. This means that if your teammates don't carry their weight then your personal mark is not as badly affected."*

Two students noted that it is also a good way to anonymously communicate how they felt about their team members. Anonymity was considered an important scaffold for second year students who were relatively inexperienced at reflecting on teamwork and exchanging feedback with their teammates. In cases where serious conflicts were detected, teams were required to hold an instructor-moderated team meeting to resolve their issues. Students were ultimately required to confront teamwork issues with their team; they could not simply remain sheltered by the anonymity of the PE system. Students reported that it was insightful to receive feedback on what their teammates thought about them and that this feedback helped them improve (11 comments). One such comment was: *"CATME allowed each person in the group to see what the group thought about them, and which aspects of teamwork they were good/bad at"*. The comments concerning feedback indicate that some students were engaging in reflection and that the CATME feedback system gave additional insight for students to understand how they performed relative to the average of their team.

Suggestions for how the peer evaluation system could be improved

Teamwork moderation and *rating scale/rubric* were the prevailing themes that emerged from students' comments on how the peer evaluation system could be improved. This cohort of Engineering students described the rubric and rating scale as ambiguous; subjective and vague; they suggested a more discrete and quantitative measurement of teamwork contribution in place of the behaviourally-anchored and descriptive rubric, which they found overly generic (5 comments). It is quite common for engineering students to demonstrate a preference for quantitative reasoning (Trevelyan, 2014) and this mindset is evident in students' suggestions; one student suggested [sic]: *"Add an estimated hours worked, even the peer evaluation is useless if the person filling it out doesn't want to be a snitch and rat out team members cause it's not worth the hassle"*. Students' comments on the lack of quantitative measurements of teamwork contribution indicates a lack of appreciation for those aspects of teamwork that are difficult to measure quantitatively (motivating the team; making sure teammates stay informed; caring that the team does outstanding work, etc). It also indicates that students tend to show disproportionate preference for technical ability and effort in labour as a measure of effective teamwork. Repeated practice of peer evaluation may be a good strategy to change engineering students' quantitative bias over time.

Some students were critical of the mark moderation process and the reliability of ratings (10 comments); one student wrote: *"My suggestion would be that the comments at the end of the section have a greater weight in the deciding how to moderate marks as there will always be some things that the questions would not be able to cover"*. Another student felt that it was harsh for a team member to be *"marked down purely because other members went above and beyond what should be expected."* and suggested that team members *"should not be scaled down simply because they may have been outperformed"*. The instructor shall take these comments into consideration for future implementations of CATME peer evaluation.

Future Work

This research is ongoing; the next iteration will report on a controlled intervention using peer evaluation and team development activities. Findings from this controlled study shall enable more insightful comparisons with the literature on the effects of PE combined with team development activities on students' development of professional skills.

Conclusions

A formal peer evaluation system and team development activities were implemented in a team-based second year Engineering design subject. Students' perceived professional skill development and their attitudes toward formal peer evaluation and team development activities were investigated with an anonymous online survey. The findings indicate that students generally recognised the benefits of both peer evaluation and team development activities, namely in making teamwork 'fairer'; holding team members accountable; improving in

teamwork KSAs; identifying strengths and weaknesses; facilitating team communications; project planning and opportunities for teamwork reflection. Following this positive response, peer evaluation has been widely implemented across the institution, with the current yearly number of unique student users estimated to be over 10,000.

References

Albanese, R., & Van Fleet, D. D. (1985). Rational behavior in groups: The free-riding tendency. *Academy of Management Review*, *10*(2), 244–255.

Bolton, M. K. (1999). The Role Of Coaching in Student Teams: A 'Just-in-Time' Approach To Learning. *Journal of Management Education*, *23*(3), 233–250.

Borrego, M., Karlin, J., McNair, L. D., & Beddoes, K. (2013). Team Effectiveness Theory from Industrial and Organizational Psychology Applied to Engineering Student Project Teams: A Research Review. *Journal of Engineering Education; Washington*, *102*(4), 472–512.

Brown, R. W. (1995). Autorating: getting individual marks from team marks and enhancing teamwork. In *Proceedings Frontiers in Education 1995 25th Annual Conference. Engineering Education for the 21st Century* (Vol. 2, p. 3c2.15-3c2.18 vol.2).

Brutus, S., & Donia, M. B. L. (2010). Improving the Effectiveness of Students in Groups With a Centralized Peer Evaluation System. *Academy of Management Learning & Education*, *9*(4), 652–662.

Chen, & Lou. (2004). Students' Perceptions of Peer Evaluation: An Expectancy Perspective. *Journal of Education for Business*, *79*(5), 275–282.

Engineers Australia. (2011). Stage 1 competency standard for professional engineer. Canberra, Australia: Engineers Australia.

Felder, R. M., & Brent, R. (2001). Effective strategies for cooperative learning. Journal of Cooperation & Collaboration in College Teaching, 10(2), 69–75.

Kaufman, D. B., Felder, R. M., & Fuller, H. (2000). Accounting for individual effort in cooperative learning teams. *Journal of Engineering Education*, *89*(2), 133–140.

Loddington, S., Pond, K., Wilkinson, N., & Willmot, P. (2009). A case study of the development of WebPA: An online peer-moderated marking tool. *British Journal of Educational Technology*, *40*(2), 329–341.

Loughran, J. J. (2002). Effective reflective practice: In search of meaning in learning about teaching. *Journal of Teacher Education*, *53*(1), 33–43.

Loughry, M. L., Ohland, M. W., & Woehr, D. J. (2014). Assessing Teamwork Skills for Assurance of Learning Using CATME Team Tools. *Journal of Marketing Education*, *36*(1), 5–19.

May, E., & Strong, D. S. (2011). Is Engineering Education Delivering What Industry Requires? *Proceedings of the Canadian Engineering Education Association*, *0*(0).

Oakley, B. A., Hanna, D. M., Kuzmyn, Z., & Felder, R. M. (2007). Best Practices Involving Teamwork in the Classroom: Results From a Survey of 6435 Engineering Student Respondents. *IEEE Transactions on Education*, *50*(3), 266–272.

Ohland, M. W., Loughry, M. L., Woehr, D. J., Bullard, L. G., Felder, R. M., Finelli, C. J., … Schmucker, D. G. (2012). The Comprehensive Assessment of Team Member Effectiveness: Development of a Behaviorally Anchored Rating Scale for Self- and Peer Evaluation. *Academy of Management Learning & Education*, *11*(4), 609–630.

Peterson, R. A., & Wilson, W. R. (1992). Measuring customer satisfaction: Fact and artifact. *Journal of the Academy of Marketing Science*, *20*(1), 61. https://doi.org/10.1007/BF02723476

Pfaff, E., & Huddleston, P. (2003). Does it matter if I hate teamwork? What impacts student attitudes toward teamwork. *Journal of Marketing Education; Boulder*, *25*(1), 37–45.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144.

Schön, D. A. (1983). *The reflective practitioner: how professionals think in action*. New York: Basic Books.

Simms, A., & Nichols, T. (2014). Social Loafing: A Review of the Literature. *Journal of Management Policy & Practice*, *15*(1), 58–67.

Smith, W. F. (2008). Twenty-one Years of the Warman Design and Build Competition. *19th Annual Conference of the Australasian Association for Engineering Education: To Industry and Beyond; Proceedings of the*, 212.

Stevens, M. J., & Campion, M. A. (1994). The Knowledge, Skill, and Ability Requirements for Teamwork: Implications for Human Resource Management. *Journal of Management*, *20*(2), 503–530.

Trevelyan, J. (2014). Flying start, no wings, wrong direction. In *The making of an expert engineer* (p. 47). CRC Press.

Willey, K., & Gardner, A. (2010). Investigating the capacity of self and peer assessment activities to engage students and promote learning. *European Journal of Engineering Education*, *35*(4), 429–443.

Willey, K., & Gardner, A. P. (2007). Building better teams at work using self and peer assessment practices. In Annual Conference of Australasian Association for Engineering Education. Australasian Association for Engineering Education.