Proceedings of the 2014 Industrial and Systems Engineering Research Conference
Y. Guan and H. Liao, eds.

Peer Evaluation and Assessment Resource (PEAR) to Assess Students' Professional Skills

Sadan Kulturel-Konak
Penn State Berks
Reading, PA USA

Abdullah Konak
Penn State Berks
Reading, PA USA

Gül E. Okudan Kremer
Penn State University Park
State College, PA USA

Ivan Esparragoza
Penn State Brandywine
Media, PA USA

Gareth Yoder
Penn State Berks
Reading, PA USA

Abstract

Professional skills expected from science, technology, engineering and math (STEM) students are well-aligned with the broad learning outcomes defined by the Accreditation Board of Engineering and Technology (ABET). While acquisition of professional skills in our graduates is increasingly crucial due to global competition and intensifying pressures on companies, the absence of a robust assessment framework limits effectiveness of pedagogical efforts by faculty. Thus, there is a need for an assessment model, which can help assess students' professional skill development across multiple disciplines. In this paper, we introduce a web-based application, called Peer Evaluation and Assessment Resource (PEAR), to assist in assessing students' development in professional skills, and we illustrate how this tool can be used for teamwork assessment. PEAR intends to streamline the processes of peer evaluations. PEAR has been designed based on the theory of Model of Domain Learning (MDL), and thus can help explain the complex interactions among knowledge, interest level, and strategies with which knowledge is gained. We illustrate the workflow of PEAR and how MDL is integrated into the peer assessment process. We also discuss the advantages of the MDL-based assessment framework compared to a traditional assessment model.

Keywords
Engineering Education, Teamwork, Self and Peer Evaluation, Assessment

1. Introduction

The Accreditation Board of Engineering and Technology (ABET)'s Criterion 3 states "an ability to function effectively as a member or leader on a technical team" as an expected learning outcomes from science, technology, engineering and mathematics (STEM) students. The competency of teamwork skills will allow STEM students to be successful in their future work environment. Therefore, it is not surprising that teamwork is ubiquitous in higher education today. In particular, undergraduate students in engineering and technology programs are involved with many forms of teamwork numerous times throughout their education. Given this setting and the expectations, an obvious question is: How effective are curricular efforts in helping students develop their teamwork skills? In order to answer this question, an effective assessment framework with accompanying instruments is needed.

Assessment of teamwork skills has been recognized to be very challenging. A primary challenge is that unlike technical skills, which can be acquired and assessed discretely, students' soft skills slowly mature throughout their education [1-3]. Another challenge is, in part, due to the fact that instructors do not know how to incorporate teamwork into the courses and thus have difficulties developing assessment tools and rubrics to evaluate it. Students are also concerned that their contributions to the team's collective output will not be fairly accounted for, or their grades will suffer by the subpar performance of underperforming team members [4]. Social loafing, i.e., the tendency of people in groups to work far less than they would individually, may be an important factor to inhibit the pedagogical benefits of teamwork [5]. Because of these reasons, teamwork is not always a positive experience for students.

The literature on the assessment of teamwork knowledge, skill, and ability (KSA) can be broadly categorized in two folds: (*i*) studies on expected KSA taxonomy development [6-12], and (*ii*) studies that aim to measure the contribution of individual students to the collective output of the team [5, 13-22]. The boundary between these two groups is not well-defined in terms of KSA as the second group's research is informed by the outcome of the first group. Self and peer evaluations have been preferred to achieve a fair summative evaluation of teamwork outcomes by determining the individual contribution of each team member, and evaluating the teamwork process and other professional skills.

This paper introduces a self and peer evaluation web-based tool, which is called Peer Evaluation and Assessment Resource (PEAR). PEAR is in its alpha-trial version. One of the challenges in peer evaluation is the labor intensive nature of data collection and analysis. Even in a small class with several teams, the number of peer ratings can easily reach into the hundreds (e.g., a team with 5 members generates 25 self and peer evaluation ratings). The sheer number of collected data prevents using peer evaluations in large classes. Often, student feedback cannot be provided in a timely manner because of the time required to summarize the collected data, diminishing its effectiveness. Confidentiality and privacy of the paper/pencil peer evaluations completed in the classroom is also a concern. PEAR addresses these administrative challenges as students can complete their self and peer evaluations through a web interface. Instructors can create self and peer evaluation rubrics, download individual student ratings and their summary reports. In addition, instructors can send feedback to students with a few clicks. Thereby, PEAR streamlines the process of self and peer evaluations, making it possible to use them in large classrooms.

There are a few online tools for administering self and peer evaluations. Self and Peer Assessment Resource Kit (SPARK) [23-25] is one of the first online tools to facilitate the administration of self and peer evaluations with the objective of measuring the individual contributions of the team members to a team project. One of the strengths of SPARK is that evaluation criteria can be customized to specific disciplines or project goals [26]. E-TAT is a work-in-progress tool proposed for measuring team effectiveness using multiple metrics such as students usage of team collaboration tools, self-evaluations, and instructor feedback [27]. The online version of the Comprehensive Assessment of Team Member Effectiveness Instrument (CATME) [8] has a set of comprehensive tools to facilitate the assessment of student contributions to teamwork through peer evaluations as well as to support faculty and students in the effective management of teams through best practices. CATME Peer Evaluation tool assists instructors to evaluate team member contributions and experiences through self and peer evaluations in five categories: (i) contribution to teamwork, (ii) interacting with teammates, (iii) keeping the team on track, (iv) expecting quality, (v) having relevant team KSA [22]. CATME Team Maker tool enables instructors to form student teams by matching student responses to a set of survey questions with the criteria defined by the instructors [28].

PEAR has similar objectives and functions to those of SPARK and CATME. Although PEAR can also be used to automate the process of self and peer evaluations by efficiently gathering and analyzing the information on individual student contributions, the primary objective of PEAR is to facilitate the assessment of student's teamwork KSA development throughout their education. To achieve this goal, PEAR utilizes an assessment framework [29] based on the Model of Domain Learning (MDL) [30, 31]. PEAR has three types of rubric items, i.e., knowledge, interest, strategic processing; and these items are mapped against the three developmental stages of the MDL, acclimation, competency, and proficiency. The details of the assessment framework of PEAR are described in [29]. This mapping allows instructors to design self and peer evaluation instruments that can appropriately tailor the assessment to the class level and/or learning objectives. Furthermore, the MDL-based assessment framework allows

instructors to make inferences about student progress enabling appropriate attributions to knowledge or interest level.

2. Advantages and Disadvantages of Self and Peer Evaluations

In self-assessment, Boud and Falchikov [32] report that higher achieving students tend to underrate their own performance relative to the instructor ratings, whereas lower achieving students overrate it. One exception to these cases is that where all students tend to overrate their own performance if the student grades depend on the ratings. Indeed, a significant concern in peer evaluation of team performance is that students inflate the achievement of their own team members due to peer pressure, or friendship. In addition, students are usually reluctant to participate in peer evaluations because they are concerned that revenge and jealousy will play a role in ratings [33].

Another concern with self and peer evaluations is psychometric reliability of student ratings. A review of the peer assessment literature by Topping [34] shows that while some instruments have satisfactory level of reliability, others have unacceptably low. Mathews [35] argues that peer evaluation results are not robust enough to assess the contribution of individuals group work. Boydell [36] reports low reliability for the ratings of more specific items. Brannick et al. [37] study the psychometric qualities of teamwork measurement metrics. In their study, the ratings of team members had lower levels of inter-rater reliability than the ratings of on-site and off-site observers', indicating that the team members were not consistent.

Despite the drawbacks mentioned above, self and peer evaluations have several advantages that make them an indispensable part of the overall assessment of teamwork. First of all, the use of peer evaluations to determine a part of the grade in a project work reduces social loafing [24, 38]. Peer evaluations can also be used as a diagnostic tool to identify and remedy dysfunctions in a team [39, 40]. More importantly, self and peer evaluations provide students with invaluable feedback about their own teamwork skills with respect to what is expected from them in real-life teams. Thereby, self and peer evaluations may promote positive changes in students' attitudes towards overall professional skills attainment.

There is also evidence in the literature that self and peer evaluations are instrumental for undergraduate students to develop their teamwork and other professional skills [25, 32]. Boyd and Cowan [41] report that self-assessment is an effective way for students to develop deep processing strategies despite the difficulties of using it.

Although some studies report weak correlations between self/peer ratings and the ratings by outside reviewers, growing body of work supports the validity of self and peer evaluations in assessing various team performance measures. McGourty et al. [42] report significant correlations between the instructor's team ratings and students' average team peer-ratings. Greguras and Robie [43] state that peer raters are more reliable than subordinates and supervisors in practical scenarios. Viswesvaran et al. [44] report acceptable levels of reliability for peer ratings. Loughry et al. [8] also report very high reliability values for the constructs used in CATME. For CATME-B, significant correlations are reported between final course grades and students' ratings [22].

3. Peer Evaluation and Assessment Resource (PEAR)

We have developed PEAR as part of a National Science Foundation (NSF) sponsored project to design an assessment framework for evaluating students' development in professional skills. The objective of the existing software tools for peer evaluation (CATME, SPARK, etc.) is to measure the individual contribution of each team member into team's collaborative output, which eventually aids to determine students' individual grade. On the other hand, PEAR aims to facilitate the assessment of students' teamwork KSA development throughout their educational journey. The assessment items included in PEAR intend to measure the attitude, behavior, or knowledge of students as perceived by their peers. According to MDL, a learner becomes an expert in a domain by going through three progressive and incremental experience-based stages, i.e., acclimation, competency and proficiency. The MDL posits that the nature of domain knowledge, strategic processing abilities, interests, as well as their interactions, are different across the acclimation, competency and proficiency stages. Within PEAR, evaluation items are categorized as knowledge, strategic processing, and interest, and are mapped against the MDL stages. Table 1 illustrates an example for the categorization and mapping of the rubric items under the "use communication technologies skillfully" competency of the teamwork communications KSA. Currently, we are in the process of content validation of rubric items by experts and validation of the mapping through field studies.

Table 1. A sample mapping for the "use communication technologies skillfully" competency/attitude

Expected Competency /Attitude (5-point Likert Rating scale)	I	K	S	A	C	P
Interested in the use of internet communications technologies to communicate	x			x		
Uses internet communications technologies skillfully		x			x	
Initiates the use of most appropriate communications technologies for specific tasks			x			x
Express a personal level interest to learn and use new internet communications technologies	x				x	
Expresses opinions and ideas in a clear manner in electronic correspondences			x		x	
Expresses opinions and ideas in a clear and concise manner with obvious connection to topic in electronic correspondences			x			x

3.1 PEAR Conceptual Model and Interface

PEAR is envisioned as an easy-to-use, flexible web-based tool. PEAR has been developed using the PHP programming language and MySQL as the backend database. Figure 1-(a) and (b) show the simplified conceptual model of PEAR and the instructor dashboard, respectively. A rubric is a collection of assessment questions that represent performance criteria or KSA. PEAR allows instructors to create custom rubrics by selecting a set of items from the assessment question repository (called master rubric) or to use prescribed rubrics in the system. The custom rubrics are owned by the instructor (e.g., IST 210 Fall 2013 in Figure 1–(b)). The main advantage of the prescribed rubrics is that their psychometric qualities will be established based on the collected data over time and will be provided to instructors. The reliability of individual items will be calculated, and the questions will be updated accordingly. PEAR can also use previously published instruments (e.g., prescribed rubrics).

An assessment is an assignment of a rubric to student teams and is used to keep track of students' ratings. Figure 1-(b) illustrates the dashboard interface when an instructor logs into PEAR. Under the assessment column, all assessments created by the instructor are displayed with their completion statuses. The rubric column displays all instructor or system created rubrics.

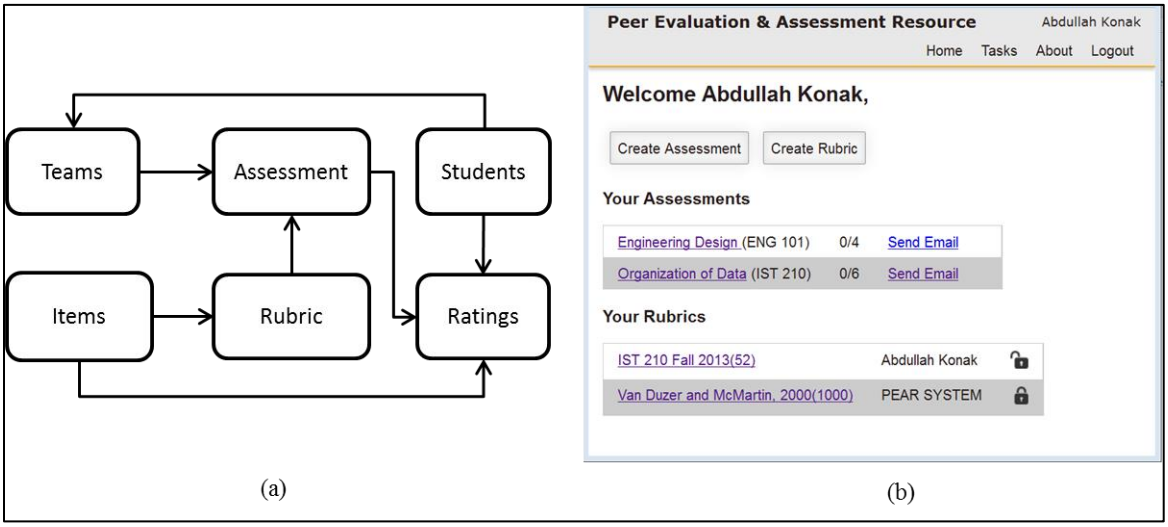

Figure 1: (a) The conceptual model of PEAR and (b) the instructor dashboard.

3.2 Assessment Creation and Report Process

In Figure 2, the process of creating a new assessment is demonstrated as follows: (a) the instructor inputs assessment related information; (b) the instructor assigns a rubric to the assessment; (c) the instructor enters the student team information as a space delimitated text list, and confirms student-to-team assignments; (d) PEAR sends each student an email that contains an encrypted individualized URL address; (e) the student clicks on the URL address and the student's rating page is loaded. The list of the students and teams can be copied and pasted from an Excel or a text file, which is readily generated by most course management systems. Students do not login to PEAR to rate their

teammates; rather when the student clicks on the URL address, the student's rating page is loaded. Figure 2-(e) illustrates an example of the student's rating page. Students are not allowed to change their ratings once they have been submitted. The instructor downloads the assessment data as a .csv file, which includes several assessment reports, including raw ratings, averaged student ratings and team ratings, breakdown of ratings based on the MDL stages and components. After an assessment is created, the same process is used to modify the assessment information. PEAR also allows the instructor to send feedback to students through email.

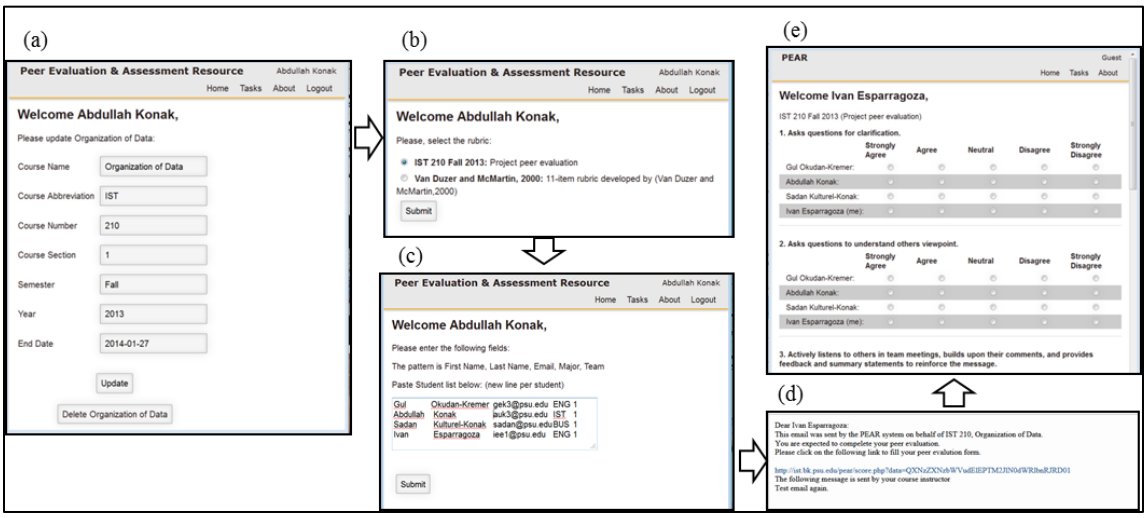

Figure 2: Screenshots of the assessment creation process

3.3 Rubric Creation Process

PEAR allows instructors to create their custom rubrics. To create a rubric, the instructor first selects the assessment area. The following screen displays all the assessment questions grouped by the expected competencies/attitudes. Then, the instructor selects a set of master rubric items and adds them to the rubric. In the rubric confirmation page, PEAR provides feedback on the distribution of the selected items with respect to the MDL components and stages. Based on this feedback, the instructor can judge the appropriateness of the assessment to the class level and/or the assessment objectives.

4. A Pilot Study

In this section, the results from an empirical study are presented to demonstrate how PEAR and the MDL-based framework can be used to assess students' teamwork KSA development. A paper-pencil instrument was designed based on the MDL-framework to assess students' KSA in the domain of teamwork communications. The survey had two sections. The first section included Likert-scale questions to measure students' interest level in the domain of teamwork communications and the perceived importance of teamwork communication skills relative to their career goals. The second section included a set of multiple-choice questions to measure student's knowledge and strategic processing abilities in the domain of teamwork communications. We adapted multiple-choice questions from the Teamwork Knowledge, Skill, and Ability (KSA) Test [6, 7]. The Teamwork KSA Test does not specify certain test items as targeting specific teamwork communications skills. We reviewed the test items to determine those that are related to teamwork communications. Next, we categorized the test items into three MDL stages (i.e., acclimation, competency, and proficiency), and into two of the MDL components: (1) knowledge, and (2) strategic processing.

In this investigation, there were 335 participants from Engineering (108), Business (140), and Information Sciences and Technology (IST) (87) programs at a land-grant university in the Northeast of United States. Students were in various stages of their undergraduate degree programs: freshman (86), sophomore (89), junior (61), and senior (99). Collected responses were analysed in two groups: freshman and sophomore (Group I) and junior and senior (Group II). Figure 3 presents the average knowledge and strategic processing scores across the MDL developmental stages as well as students' interest levels. This output is an example for reports that PEAR can generate. Based on the empirical results presented in Figure 3, it could be concluded that students are still at the acclimation stage because

the scores at the competency and proficiency levels are quite low. Students' perceptions of the importance of teamwork communication skills are high in both groups as seen in Figure 3, but they are not ready yet to commit individual effort for the mastery of those skills as indicated by the low individual interest scores. In both knowledge and strategic processing components, the average score decreases significantly from the acclimation stage to the proficiency stage as expected.

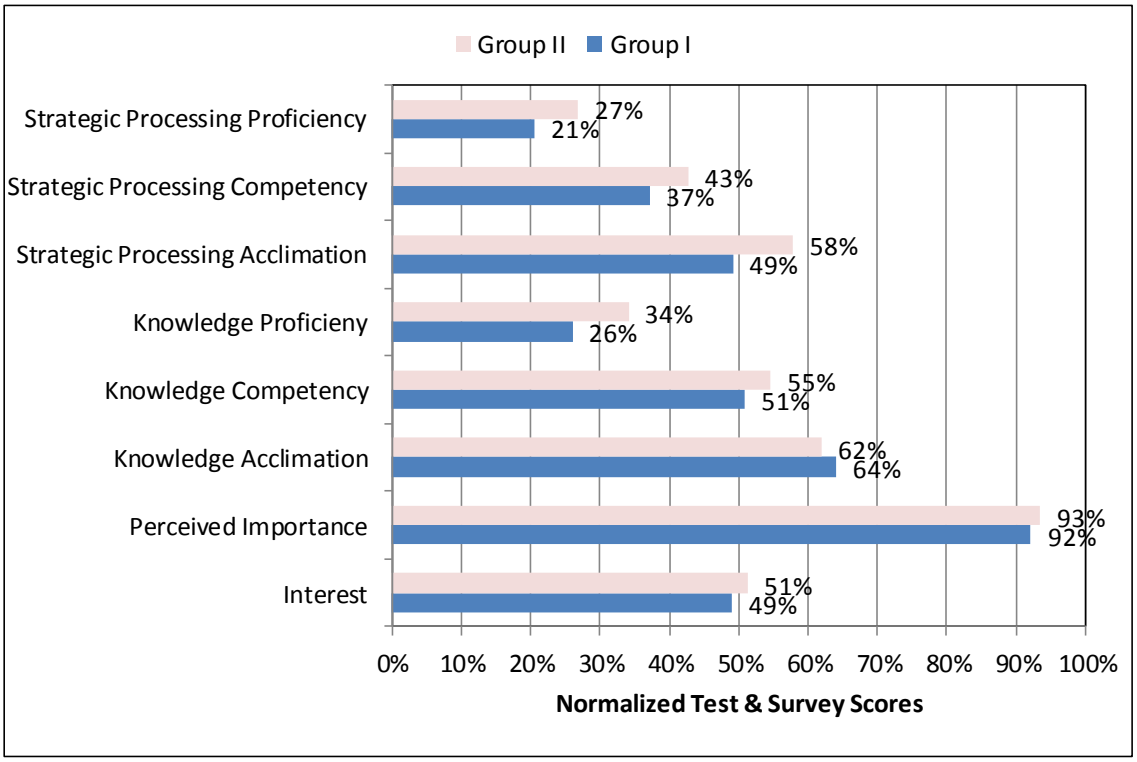

Figure 3: Knowledge and strategic processing scores across MDL stages

5. Conclusions and Further Work

In engineering education, team projects and collaborative learning strategies are increasingly being used to prepare students for today's engineering careers which require functioning in multidisciplinary teams for success. Assessing students' team work skills is a difficult, but a necessary task to give them guidance for personal development and to design effective interventions to address weaknesses. Self and peer evaluations are frequently used to evaluate individual student performance in teamwork and assess student's teamwork knowledge, skills, and abilities. Peer Evaluation and Assessment Resource (PEAR) is a promising online tool to facilitate the self and peer evaluation process. Currently, PEAR is in its alpha testing phase. After a rigorous beta testing phase, data will be collected to evaluate the reliability of PEAR instruments and the validity of the proposed MDL-based framework.

Acknowledgements

This work is supported by the National Science Foundation (NSF) grant # DUE-1141001.

References

1. Perry, W.G., 1970, "Forms of intellectual and ethical development," New York: Rinehart and Winston Inc.
2. King, P.M. and Kitchener, K.S., 1994, Developing Reflective Judgment: Understanding and Promoting Intellectual Growth and Critical Thinking in Adolescents and Adults. Jossey-Bass Higher and Adult Education Series and Jossey-Bass Social and Behavioral Science Series, Jossey-Bass San Francisco, CA.
3. Alexander, P.A., Murphy, P.K., Woods, B.S., Duhon, K.E., and Parker, D., 1997, "College instruction and concomitant changes in students' knowledge, interest, and strategy use: A study of domain learning," Contemporary Educational Psychology, 22(2), 125-146.

4. Willcoxson, L.E., 2006, "It's not Fair!": Assessing the Dynamics and Resourcing of Teamwork," Journal of Management Education, 30(6), 798-808.
5. Aggarwal, P. and O'Brien, C.L., 2008, "Social loafing on group projects structural antecedents and effect on student satisfaction," Journal of Marketing Education, 30(3), 255-264.
6. Stevens, M.J. and Campion, M.A., 1994, "The knowledge, skill, and ability requirements for teamwork: Implications for human resource management," Journal of management, 20(2), 503-530.
7. Stevens, M.J. and Campion, M.A., 1999, "Staffing work teams: Development and validation of a selection test for teamwork settings," Journal of Management, 25(2), 207-228.
8. Loughry, M.L., Ohland, M.W., and Moore, D.D., 2007, "Development of a theory-based assessment of team member effectiveness," Educational and Psychological Measurement, 67(3), 505-524.
9. Hackman, J.R., 1980, "Work redesign and motivation," Professional Psychology, 11(3), 445-455.
10. Hackman, J.R., 1983, A normative model of work team effectiveness. Office of Naval Research: Arlington, VA.
11. Wageman, R., Hackman, J.R., and Lehman, E., 2005, "Team diagnostic survey: Development of an instrument," The Journal of Applied Behavioral Science, 41(4), 373.
12. Baker, D.P. and Salas, E., 1992, "Principles for measuring teamwork skills," Human Factors: The Journal of the Human Factors and Ergonomics Society, 34(4), 469-475.
13. Matsuura, S., 2006, "An evaluation method of project based learning on software development experiment," SIGCSE Bulletin, 38(1), 163-167.
14. Cooper, G. and Heinze, A., 2007, "Centralisation of assessment: Meeting the challenges of multi-year team projects in information systems education," Journal of Information Systems Education, 18(3), 345-356.
15. Smith, L., 2007, "Teaching the Intangibles," T + D, 61(10), 23.
16. Smith, H., Smarkusky, D., and Corrigall, E., 2008, "Defining Projects to Integrate Evolving Team Fundamentals and Project Management Skills," Journal of Information Systems Education, 19(1), 99.
17. Ge, Y., Han, P., and Xu, B., 2009, "A comprehensive approach for software group project assessment based on students' evaluation," in 4th International Conference on Computer Science Education (ICCSE 2009). Piscataway, NJ, USA 1553-8.
18. Clark, N., Davies, P., and Skeers, R., 2005, "Self and peer assessment in software engineering projects," in The 7th Australasian conference on Computing education. Darlinghurst, Australia 91-100.
19. LeJeune, N., 2006, "Assessment of individuals on CS group projects," Journal of Computing Sciences in Colleges, 22(1), 237.
20. Loughry, M.L., Ohland, M.W., and Moore, D., 2006, "Behaviorally Anchored Peer Evaluation of Team Member Effectiveness," in Proceedings of the Annual Meeting of the Southern Management Association.
21. Van Duzer, E. and McMartin, F., 2000, "Methods to improve the validity and sensitivity of a self/peer assessment instrument," Education, IEEE Transactions on, 43(2), 153-158.
22. Ohland, M.W., Loughry, M.L., Woehr, D.J., Bullard, L.G., Felder, R.M., Finelli, C.J., Layton, R.A., Pomeranz, H.R., and 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.
23. Freeman, M. and McKenzie, J., 2002, "SPARK, a confidential web–based template for self and peer assessment of student teamwork: benefits of evaluating across different subjects," British Journal of Educational Technology, 33(5), 551-569.
24. Willey, K. and Freeman, M., 2006, "Improving teamwork and engagement: the case for self and peer assessment," Australasian Journal of Engineering Education, 12 2006-2002.
25. Willey, K. and Freeman, M., 2006, "Completing the learning cycle: The role of formative feedback when using self and peer assessment to improve teamwork and engagement," in Proceedings of the 17th Annual Conference of the Australasian Association for Engineering Education: Creativity, Challenge, Change; Partnerships in Engineering Education. 751.
26. Willey, K. and 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.
27. Petkovic, D., Thompson, G., Todtenhoefer, R., Huang, S., Levine, B., Parab, S., Singh, G., Soni, R., and Shrestha, S., 2010, "Work in progress, e-TAT: Online tool for teamwork and soft skills; assessment in software engineering education," in Frontiers in Education Conference (FIE), 2010 IEEE. S1G-1-S1G-3.
28. Layton, R.A., Loughry, M.L., Ohland, M.W., and Ricco, G.D., 2010, "Design and validation of a web-based system for assigning members to teams using instructor-specified criteria," Advances in Engineering Education, 2(1), 1-9.

29. Kulturel-Konak, S., Konak, A., Esparragoza, I.E., and Kremer, G.E.O., 2013, "Assessing professional skills in STEM disciplines," in IEEE Integrated STEM Education Conference (ISEC). Princeton, NJ 1-4.
30. Alexander, P.A., 2003, "The development of expertise: The journey from acclimation to proficiency," Educational Researcher, 32(8), 10-14.
31. Alexander, P.A., Sperl, C.T., Buehl, M.M., Fives, H., and Chiu, S., 2004, "Modeling domain learning: Profiles from the field of special education," Journal of educational psychology, 96 545-557.
32. Boud, D. and Falchikov, N., 2007, Rethinking assessment in higher education: Learning for the longer term, Routledge
33. Taggar, S. and Brown, T.C., 2006, "Interpersonal affect and peer rating bias in teams," Small group research, 37(1), 86-111.
34. Topping, K., 1998, "Peer assessment between students in colleges and universities," Review of Educational Research, 68(3), 249-276.
35. Mathews, B.P., 1994, "Assessing individual contributions: Experience of peer evaluation in major group projects," British Journal of Educational Technology, 25(1), 19-28.
36. Boydell, D., 1994, "The use of peer group review in the assessment of project work in Higher Education," Mentoring & Tutoring: Partnership in Learning, 2(2), 45-52.
37. Barrick, M.R., Stewart, G.L., Neubert, M.J., and Mount, M.K., 1998, "Relating member ability and personality to work-team processes and team effectiveness," Journal of applied psychology, 83(3), 377.
38. Brooks, C.M. and Ammons, J.L., 2003, "Free riding in group projects and the effects of timing, frequency, and specificity of criteria in peer assessments," The Journal of Education for Business, 78(5), 268-272.
39. Dyrud, M.A., 2001, "Group projects and peer review," Business Communication Quarterly, 64(4), 106-111.
40. Smith, H.H. and Smarkusky, D.L., 2005, "Competency matrices for peer assessment of individuals in team projects," in Proceedings of the 6th conference on Information technology education. Newark, NJ, USA. 155-162.
41. Boyd, H. and Cowan, J., 1985, "A case for self-assessment based on recent studies of student learning," Assessment and Evaluation in Higher Education, 10(3), 225-235.
42. McGourty, J., Dominick, P., and Reilly, R.R., 1998, "Incorporating student peer review and feedback into the assessment process," in Frontiers in Education Conference, 1998. FIE'98. 28th Annual. 14-18.
43. Greguras, G.J. and Robie, C., 1998, "A new look at within-source interrater reliability of 360-degree feedback ratings," Journal of Applied Psychology, 83(6), 960-968.
44. Viswesvaran, C., Schmidt, F.L., and Ones, D.S., 2005, "Is there a general factor in ratings of job performance? A meta-analytic framework for disentangling substantive and error influences," Journal of Applied Psychology, 90(1), 108-131.