International Journal of Engineering Education Vol. 27, No. 2, pp. 310–322, 2011
Printed in Great Britain

0949-149X/91 $3.00+0.00

Peer Assessment for Undergraduate Teamwork Projects in Petroleum Engineering*

Y. CINAR[1], A. BILGIN[2]
[1] School of Petroleum Engineering, Faculty of Engineering, University of New South Wales, Australia.
E-mail: y.cinar@unsw.edu.au
[2] Department of Statistics, Faculty of Science, Macquarie University, Australia.
E-mail: ayse.bilgin@mq.edu.au

This paper presents an analysis of the peer assessment data used to derive individual marks for teamwork projects assigned in four petroleum engineering courses. The collected data represents a wide range of groups composed of international and local students, male and female students. The analysis emphasises peer assessment as a viable tool for evaluating individual contributions, as long as human bias is adequately minimised for fair assessments. We applied the normalisation factor technique to address biases. The study demonstrates that poorly performing students in the groups can be identified by utilising peer assessment. However, to be successful additional judgement is required from the lecturer, based on the observation of team dynamics. It was determined that a significant number of students (about 40%) gave the highest marks to peers for all assessment criteria; in particular a high proportion of the international students in our study were more generous than the local peers. In four groups out of 22 groups, all team members gave the highest rates to each other for all criteria thus severely undermining the peer assessment process. All these observations underline the necessity to give students adequate training in the method when using peer assessment; and care needs to be taken when groups are formed to enable effective group dynamics.

Keywords: peer assessment; teamwork; petroleum engineering; learning styles; international students

1. Introduction

Today's engineers no longer do their work alone. The new engineer must know the types of skill that each profession brings to the table and understand the role that each plays as a member of a team [1]. There is an increasing demand from employers that graduates should have a range of transferable skills in addition to their discipline-specific knowledge and understanding. The transferable skills are defined by accreditation organisations as those that include teamwork, problem solving, communication, lifelong learning, time management, interpersonal relationship skills and initiative, as well as numeracy and IT literacy. In a 2003 survey undertaken by The World Chemical Engineering Council, covering 2,158 students from 63 countries, teamwork was ranked first in relevance to work [2].

The oil and gas industry is no exception. The engineers in this industry are always part of a resource management team consisting of geoscientists, a variety of engineers and other professionals (often from many countries). The importance of teamwork in petroleum engineering education was first mentioned in the early 1990s when a series of colloquia on petroleum engineering education were initiated by the Society of Petroleum Engineers [3]. The colloquia concluded that, because of the complexity of the upstream process, teamwork in cross-disciplinary and multicultural environments is an essential skill. In the 2000 colloquium [4], the desired petroleum engineer graduate was described as a well-rounded, broadly-based, adaptable and motivated problem solver with good communication, interpersonal and team skills.

Teamwork is an increasingly popular component of undergraduate curricula in order to meet the needs of industry. Teamwork projects adapted in the undergraduate curricula have the potential to improve the student's skill for problem solving, communication through oral presentations and written reports, time and project management, interpersonal relationships and technical knowledge. Such projects also help to meet the goal of active engagement in the learning process. When students work together in groups to solve complex and authentic problems, it can assist them in developing not only content knowledge but additionally generic graduate attributes, such as problem-solving, reasoning and communication skills. Moreover, peer learning takes place, during which students learn from each other while they are working together and assessing their own and each other's performance [5].

Naturally there are both merits and demerits of teamwork in higher education. Dyball et al. [6]

* Accepted 22 January 2011.

suggest, drawing from education, management and psychology literature, that benefits for the students include:

(1) the opportunity to work on more comprehensive assignments, thus extending learning;
(2) gaining an insight into group dynamics and processes;
(3) the development of interpersonal skills;
(4) exposure to the viewpoints of others;
(5) preparing them for the 'real viewpoint';
(6) promoting reflection and discussion.

An additional advantage is that teamwork skills are viewed by academics and employers as being the most important generic attributes that students need to acquire to prepare for the world of work [7, 8]. On the other hand, Walker [9] and Spalding et al. [10] have pointed out potential disadvantages of teamwork, such as:

(1) interpersonal conflict within the team leading to an inability to complete the task;
(2) unequal distribution of the workload;
(3) the possibility of reduction of mental effort through shirking tasks, thus debilitating learning;
(4) students' negative reactions to group learning experiences, for example, when they believe the teacher has poor team skills or does not help the team.

2. Peer assessment

2.1 Peer assessment for individual marks

The use of group projects for learning and assessment is a way of teaching engineering students to work in teams, as this reflects the situation in the workplace. The usual practice for assessment is to assign one overall mark to the group, but this can be problematic if it is necessary to assign grades on an individual basis. Since team members are in a much better position to judge individual contributions than a tutor or lecturer, it has been proposed that this function could be performed by the team members themselves [11, 12, 13]—that is, peer assessment. This also mirrors the workplace, where equal status people often comment on their colleagues' work and provide feedback to each other for improvement. By involving students in the process of assessing their own learning (self-assessment) and/or giving them the responsibility to judge the outcomes of their own and others' learning, they can practice much-needed skills for the workplace.

Peer assessment in higher education has been the focus of much investigation, both in terms of its process [14] and its outcomes [15]. Identified benefits include:

- increased student responsibility and autonomy;
- student insight into assessment procedures;
- harder working students;
- opportunities for increased levels of feedback;
- a context that encourages deep learning [16].

Furthermore, students may develop the ability to work cooperatively, reflect upon and be critical of others' work and receive critical appraisals of their own work [17]. Their active involvement in the assessment experience promotes the acquisition of lifelong skills, thus enabling them to learn on their own after graduation [18, 19, 20].

Evaluation of individual contributions can be made against a list of categories [21, 22, 23], or be based on a single overall score (a holistic approach). Lejk and Wyvill [24] argue that the holistic approach leads to higher inter-rater agreement and is thus better suited for the purposes of summative assessment, though the category-based approach may be more useful as formative assessment.

Concerns have nevertheless been expressed about the validity and reliability of peer assessment. Cassidy [15], for example, suggests that reliability for grading purposes may be uncertain since students may not have the skills or understanding to make judgements about the work of others, or to provide useful feedback. There is also a perception in the academic community that there is a lack of objectivity in the peer assessment process [25], and specific biases in peer assessment like gender differences and over-generous marking of friends have been raised as a potential difficulty [16].

2.2 Ensuring fairness

There are a number of systems for peer assessment, such as having a team agree on the contribution of individual members, or having the members submit their own evaluations anonymously. But how then do you weight these assessments? Various methods have been suggested; for example, these grades might form a fixed fraction of the individual students' overall grade with the other fraction coming from the team-performance grade; or the peer assessment might be used to develop a multiplying factor for each student that is then used to calculate the individual grades from the team performance grade.

For peer assessments that consist of scores submitted individually by team members, it is essential to ensure that these evaluations are used in a fair and equitable way to distribute marks to individual members of the team: this is essentially a peer-assessment moderation problem [23, 26]. A major problem is that students individually use different standards to adjust or evaluate their group members [23] or, in other words, the evaluation is not fully objective: students are inevitably subjective during the peer assessment process. A number of studies have addressed this problem in an attempt to increase objectivity and counter biases [27, 23, 26].

Lejk and Wyvill [28], for instance, have recommended excluding self- assessment from the process as a basic measure to try and improve the objectivity of the grading. They observe that objectivity can be weakened by such factors as the tendency of high-performing students to underrate their own performance (as compared to their peers' assessments), while low-performing students tend to overrate themselves; this can be quite unconscious but can also result from a deliberate attempt to inflate the student's own grade. Goldfinch [29] suggested that even without self-assessment, some of the students may deliberately give lower rates to their peers in order to score higher marks. She proposed a method of peer assessment—without self assessment—to derive individual marks (originally proposed in [21]; and further developed in [29]). Li [23] then modified this grading system specifically to address potential bias problems and developed a 'bias factor' technique; and Bushell [26] proposed further modifications to the method (we describe Li's technique in more detail in Section 3.2).

Despite the inherent problems, Sluijsmans et al. [30] have concluded that peer assessment is a skill that can and should be learned through focused training. Bilgin and Fraser [31] emphasise that an essential ingredient for its success is that guidance is available with regard to the criteria and standards that will inform the students' decisions. It is therefore important to provide students with information about the importance of peer marking [32].

We instituted a programme of group projects in petroleum engineering courses incorporating peer assessment to derive individual marks from a group mark, with a focus on using techniques to correct bias. A significant feature of our university context is the demographic composition of our classes and we were interested in whether this would affect the peer assessment process. Half our students were from other countries and represented a wide diversity of learning and cultural backgrounds. It has been noted by many researchers that changing between learning environments can lead to some difficulties in adjusting to the new context due to differences in learning styles [33–37]. For instance, for those who are accustomed to a top-down, teacher-driven learning environment, the whole concept of peer assessment may prove quite difficult to comprehend.

In this paper, we present an analysis of our results for four petroleum engineering courses across two years. We followed the technique proposed by Li [23] to minimise variations in student approaches: we examined the individual marks before and after the normalisation process, and attempted to identify any biases that were operating. Essentially we were testing Li's method in order to derive an accurate individual grade for each of the students. In conclusion, we offer some recommendations about how to conduct peer assessment successfully and raise some questions that may warrant further research.

3. The study

3.1 Description of the project

We instituted a teamwork study within a petroleum engineering programme offered at a university in Sydney, Australia, in four petroleum engineering courses. These courses were offered in Year 3 (junior) and Year 4 (senior); Table 1 shows a summary of the characteristics of the courses. Team projects were assigned to groups of 4–7 students, and the students were allowed to sign into any group they wanted. The teams were asked to search for an open-ended physical problem, to define and solve the problem, write a technical report and make an oral presentation. Five weeks were given to the teams to complete the project. Each team was

Table 1. Description of enrolments in all courses

Course #	Total enrolment	Local %		International %		Teams
		Male	Female	Male	Female	
1—Year 3	23	48	4	35	13	1 team of 5 students 3 teams of 6 students
2—Year 3	44	39	2	39	20	3 teams of 7 students 3 teams of 6 students 1 team of 5 students
3—Year 4	25	64	0	24	12	5 teams of 5 students
4—Year 4	29	48	3	34	14	5 teams of 5 students 1 team of 4 students
Total	121	48	2	34	16	

required to hold a weekly meeting and take minutes of this meeting, and there were also weekly meetings of the team with the lecturer.

The lecturer assessed teamwork dynamics based on the minutes of the team meetings and on the joint meetings. The project was allocated an assessment value of 30% of the total mark for the course. Of this, the teamwork dynamics, technical report, and oral presentation were used to determine a total mark for each group, which represented a maximum of 30% of each individual's mark for the project). Peer assessment marks were then used to determine the overall individual marks.

There were 121 students, 83% of whom were male and 17% female. Half the students enrolled were international—the overseas students mostly came from countries in the Asia Pacific region (78%), while the rest were from the Middle East (10%), Europe (7%), Africa (3%) and North America (2%). There were 22 teams in total, three of which were local students only, five were international students only and 14 teams were blended local and international; nine groups were male only, while the rest were composed of male and female students.

After completion of the assignment, each student filled in a peer assessment form (Table 2) anonymously to rate each of their peers' contributions to the group project on a scale from 1 to 5 (1-poor, 5-outstanding). This form was designed in relation to logistics, leadership, teamwork dynamics, intellectual and research/writing/editing, since these characteristics of the project were considered to be important elements of learning in the courses studied. We were following Li's advice [23] that each form should be specifically designed for each group project and that 'care must be taken to formulate the peer assessment form [as] the effectiveness of the entire assessment procedure is anchored on this form' (p. 11); any methods designed to correct human bias during the process cannot cure any problems caused by the design of the assessment form.

Students were additionally allowed to submit written feedback on their peers, which helped us to understand the team dynamics and also to solve a few disputes within the teams. Self-assessment was also included in the questionnaire to allow us to observe students' self-rating, but it was not used in grading the individual performances.

3.2 Application of methods for correcting bias

We have discussed earlier the problems that arise when attempting to assign individual grades to students for participation in teamwork. The simplest method of grading individuals entails summing up an individual effort rating (IER) and an average effort rating from the raw ratings given to each team member by the team. An individual weighting factor (IWF) is then calculated by dividing the IER by the average effort rating. An example worksheet for one of the cases analysed in this study is presented in Appendix A. The process is relatively simple, and a student who is rated by peers as average should have a weighting factor of 1.

A few inherent weaknesses in this simple approach have been identified [23]. For example, in the sample given in Appendix A, the IWF varies from 0.95 to 1.04 about a mean of 1 and a standard deviation of 0.04, meaning that the deviation is small and there was close agreement between team members in their ratings. However, these deviations varied and sometimes they were much larger than the example here. Appendix B shows another example with a relatively high deviation. Both examples are real cases and according to the observations the instructor made, the team in Appendix A performed well together whereas the team in Appendix B had experienced poor group dynamics. The group marks given by the lecturer also support this observation. In the poorly performing team, students 1 and 5 did not participate in the team activities as much as the others: this unequal distribution of workload is clearly shown in the raw marks given by the other four team members to both students for their contributions. It is interesting to note that both these students marked each other quite generously.

One of the issues in peer assessment that causes bias is that a student could be much more generous to peers than the rest of the group (like student 1 in

Table 2. The peer assessment form

Rubrics	Questions	1	2	3	4	5
Logistics	Did s/he participate in group meetings in an organised fashion & meet group deadlines?	❍	❍	❍	❍	❍
Leadership	Did s/he provide leadership through listening to others and helping the group to function as a team?	❍	❍	❍	❍	❍
Teamwork	Did the person share group responsibility without argument or disruption?	❍	❍	❍	❍	❍
Intellectual	Did s/he provide useful ideas, helpful suggestions and feedback to the group?	❍	❍	❍	❍	❍
Research/ Writing	Did s/he help in researching and writing the final paper in accordance with the group decisions?	❍	❍	❍	❍	❍

Appendix B). On the other hand, some students may be hard markers in comparison to other members (like student 1 in Appendix A). Li [23] proposed a bias factor in order to prevent this kind of unfair biasing of the peer assessments. The bias factor is defined as the ratio of the rating given by each team member to others, in relation to the average effort rating. The value for the nominator is obtained simply by summing up all the ratings given to other members by the particular member. A student is unbiased if s/he has a bias factor of 1, while a bias factor of less than 1 means that the student tends to rate the rest of the group relatively low. It is suggested that, for very high/low bias factors, the teacher's observations of the group must also be consulted.

Li [23] also introduced a normalisation factor—this is the inverse value of the bias factor—to adjust the marks for extreme cases, which entails multiplying the raw marks by the normalisation factor. Appendices A and B show sample calculations. In essence, human subjectivity cannot be removed entirely, but by applying the normalisation factor it can be averaged out for the whole group. It is clear from the procedure that the normalisation becomes redundant when there is no significant bias. In Appendix A, student 1 created the largest bias because of hard marking compared to all the other team members. As a result, the normalisation yields the largest change in that student's mark. Similarly, in Appendix B, student 1 and student 5 created significant biases mainly because the other team members all gave them the lowest marks. The reason why student 5 received the largest change in mark is that student 1 was more generous than student 5.

Applying the normalisation factor, however, can cause changes in ranking between the best-performing students. For instance, it can be seen in the sample sheets in Appendix A and B that the ranking changed. As a test of this, it can be seen that in both sample sheets the ranking changed. In Appendix A, the bias created by student 1 was corrected by using Li's normalisation method at the expense of his ranking, which dropped from 2 to 3; the rest of the student rankings in Appendix A did not change. In Appendix B, two poorly performing students (student 1 and student 5) exchanged their ranks because student 1 was too generous in comparison to student 5; this bias was corrected. The rest of the ranking did not change in this case either. Similar changes in ranking were observed in other 20 peer assessments. In five of them, the best-performing student was ranked down after the normalisation process. The main reason for this was that the best performer gave lower rates to the second or third-best performing student. After the normalisation process, that student lost ranking.

Another issue with peer assessment is that the best-performing students from poorly performing teams can obtain higher individual marks than the best-performing students from the best-performing teams. Looking at the samples given in Appendices A and B, student 2 and student 6 in Appendix B received higher marks than all members in Appendix A, even though the team in Appendix B received an almost 20% lower mark than the team in Appendix A.

These kinds of cases need an additional judgement by the lecturer based on the observation of the group dynamics as well as on the performance of individual students in other parts of the course. One of the practical solutions for this problem is to set a maximum increase in individual marks for all groups.

4. Results and discussion

In this section we discuss our overall results and the difficulties that arose in our attempt to make adjustments for bias.

Appendix C shows a summary of the characteristics of the groups. The group marks are those given by the lecturer to the groups (out of 30 marks). The averages and standard deviations of the group marks are 23.8 and 2.5 for Year 3 groups (Courses 1 and 2) and 24.8 and 2.7 for Year 4 groups and 24.3 and 2.6 overall.

As regards the composition of the groups, there were five groups of international students only, three groups of local students only, and 14 groups of a mixture of local and international students. There were nine groups of male students only and the rest consisted of both male and female students. There were four groups who gave the maximum rates to each other in the peer assessment, i.e. zero bias and standard deviation of IWF. It is interesting to note that three of these groups were those composed of international students only. An analysis of individual rates shows that 53 students gave the highest rates possible to their peers, 72% of which were international students: that is, the international students were more generous than the local peers. Giving everyone the same highest mark obviously undermines the whole purpose of peer assessment and makes the bias in the assessments even more severe.

Figure 1 shows the relations between the group mark, standard deviation of IWFs and the difference between the lowest and highest bias factors. There is a significant positive correlation between the standard deviations and the differences in the bias factors in the groups (Spearman's rho=0.883, $p<0.001$). Although there are negative trends between group marks and bias differences as well as

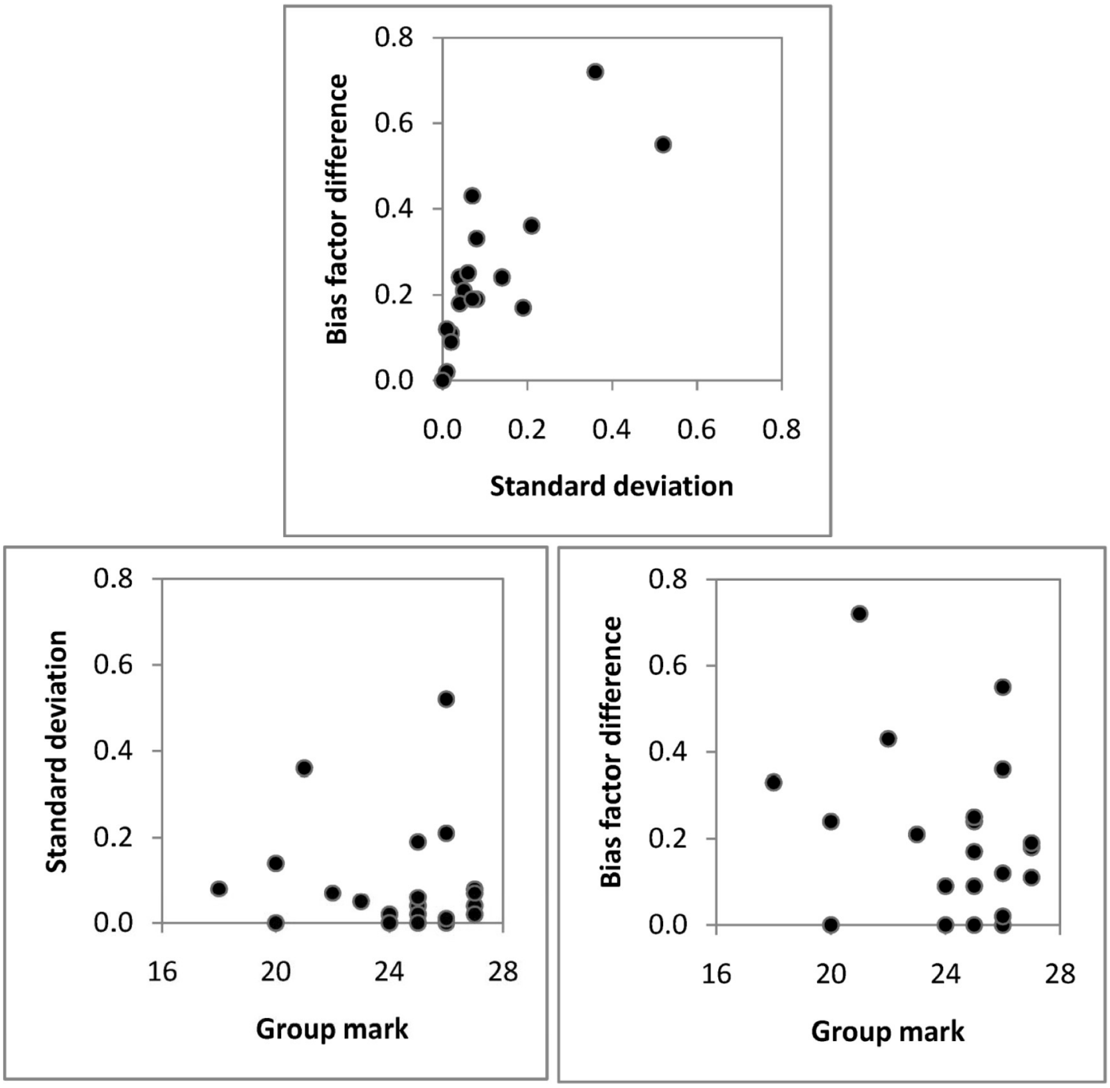

Fig. 1. Distributions of bias factors and standard deviations of IWFs for all groups.

group marks and standard deviations of IWFs, these are not significant. Nevertheless, Fig. 1 shows that when group marks increase, both bias differences and standard deviations of IWFs decrease.

Table 3 shows a comparison between the groups of differing compositions. The results show that groups of only local students obtained higher group marks—assigned by the lecturer—than the other group compositions. Since the teamwork marks were based on the oral presentations and written reports, this difference may be explained by differences in language skills. On the other hand, the blended groups formed of local and international students obtained lower marks compared to the other types of groups, with a higher standard deviation of IWFs and more significant bias in the peer assessment. This might be due to cultural differences between students preventing groups from acting as a unified entity (i.e. unstable group dynamics). The analysis also shows that the gender composition in the groups yields different biases, with less bias in the blended groups and more bias in the male-only groups.

Figure 2 shows the impacts of both the percentages of international students as well as female students in the blended groups on the group average marks, standard deviations of IWFs and the bias factors. It seems that neither the percentage of international nor the percentage of female students had an impact on the overall group marks. On the other hand, Fig. 2 suggests that when the percentage of international students or the percentage of female students in a group increases, then the differences in the bias factors decrease. This could be for a number of reasons. May be cultural differences; for instance, students from some cultures may not be comfortable with criticising others. We demonstrated before that the international students gave rates to peers more generously than the local students did. It is important to note that the majority of the female

Table 3. Comparison of average group marks, standard deviations and bias factors

Groups of	Avg. group mark	Avg. SD of IWFs	Avg. bias difference
international students only	24.6	0.05	0.22
local students only	26.0	0.05	0.17
international and local students	23.9	0.13	0.25
male students only	24.5	0.13	0.23
male and female students	24.4	0.06	0.16

students were international, and here too it is possible they may have had difficulty making criticisms. Figure 3 (comparing international students' ratings and the raw rates) shows similar results, although the trends are not as clear as in Fig. 2.

Figure 4 (comparing female students' ratings and the raw rates) displays more obvious trends for the percentage of female students in the blended groups; when the percentage of female students in a group increased, then the number of students who gave full rates to each other also increased. Because the majority of the female students were international,

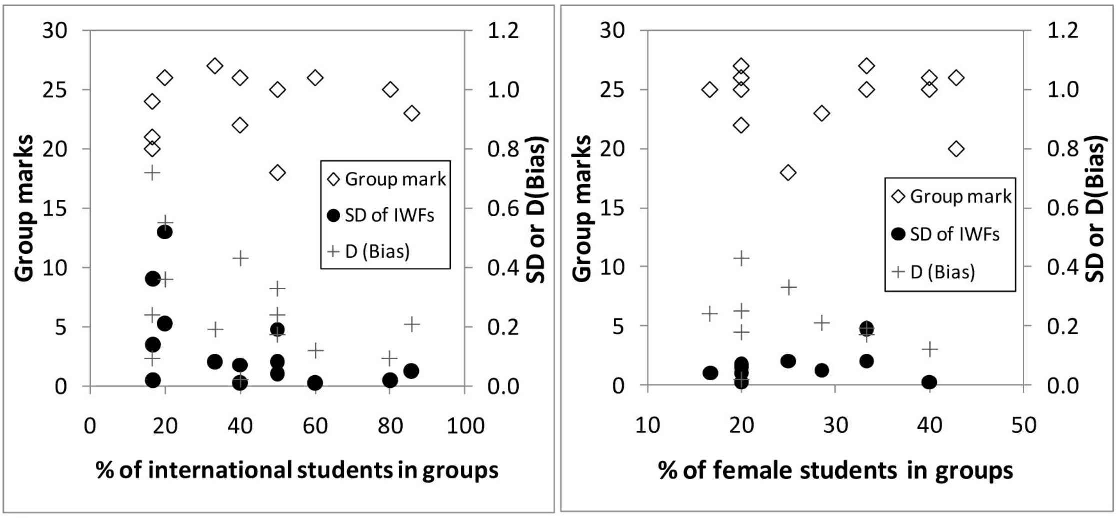

Fig. 2. Impacts of the percentages of international and female students in blended groups on group marks, standard deviations of IWFs and the bias factors.

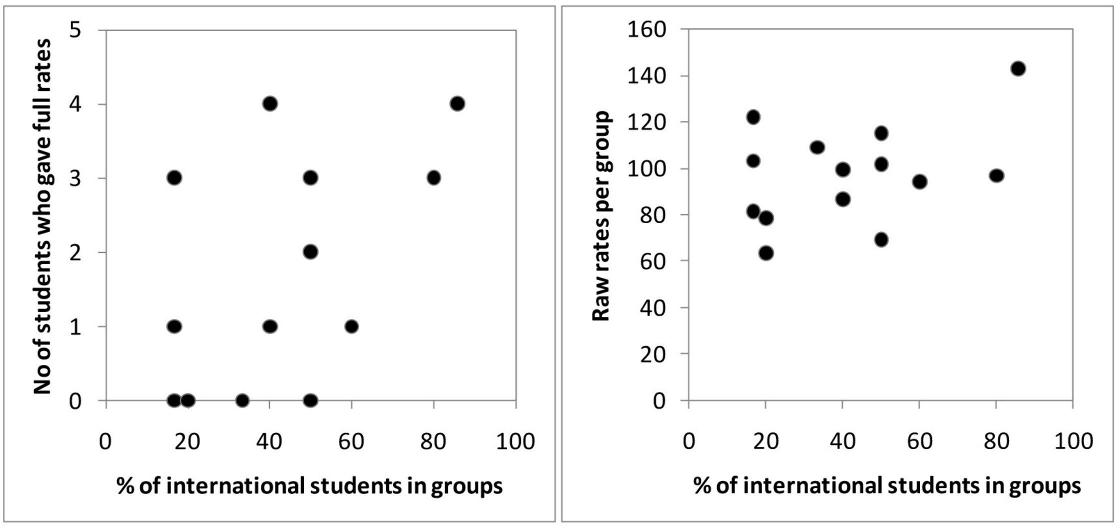

Fig. 3. Variations of no. of students who gave full marks in blended groups and raw rates per group as function of percentage of international students in blended groups.

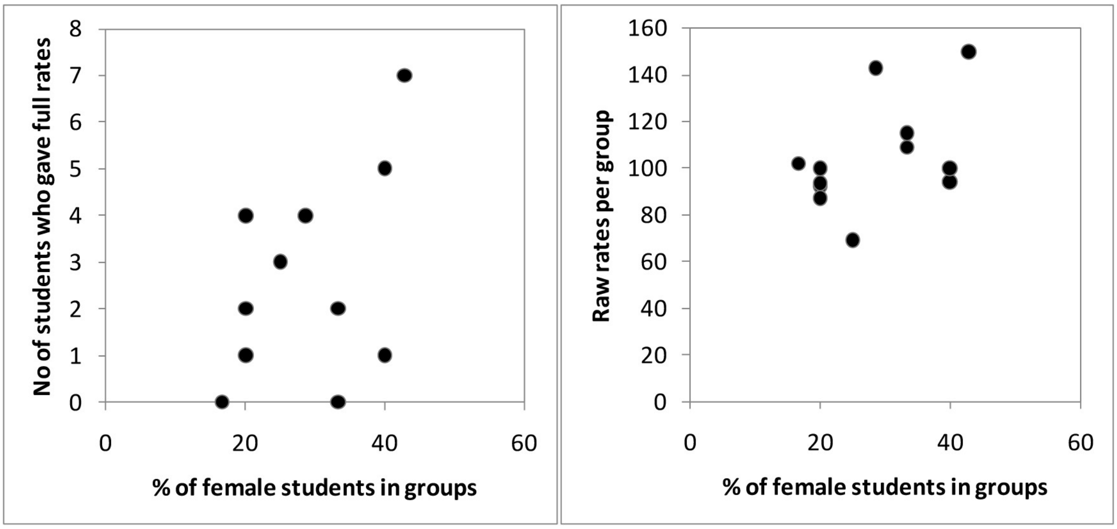

Fig. 4. Variations of no. of students who gave full marks in blended groups and raw rates per group as a function of percentage of female students in blended groups.

it is difficult to determine whether this result is due to gender bias or not.

Perhaps our most interesting finding was that the international students seemed to be more generous than the local peers in their assessments. We derive this finding from the data collected (see Fig. 3). We believe that the main cause for this finding is probably that the majority of international students were from countries of the Asia Pacific and Middle East regions whose education systems are different from the Australian system. Most likely, peer assessment is not widely practised in these systems, so students had no prior experience with peer assessment. We suspect this finding would not be limited to the discipline of petroleum engineering.

There are several potential solutions to tackle this problem. First of all, there exists a clear need to train the students for the merits of peer assessment in teaching and learning as well as today's professional life. Proper design of the assessment form is essential to ensure fair judgement by the students. A possible way of achieving this is by involving students in the design process, to enable them to understand the meanings of the terms on the form and the expectations of the lecturer about performance.

One way to offset the bias due to cultural differences may be to ensure that team compositions are as mixed as possible. Blending the groups may also help to overcome any potential gender bias (because we found that groups comprised of males only received lower group marks compared to mixed groups).

One of the advantages of peer assessment is that poorly performing students can be identified and their marks suitably adjusted. However, this also requires a judgement by the lecturer based on their direct observation of team members' individual performance. In our programme we based this judgment on the weekly meetings between the lecturer and the team, and minutes taken during the team meetings.

5. Concluding remarks

The analysis leads us to the following conclusions:

(1) Peer assessment is a viable tool for evaluating the individual contributions to a discipline-specific teamwork project; however, the human bias must be reduced in the assessments.
(2) The normalisation process proposed by Li [23] does reduce the bias.
(3) International students seemed to be more generous than local peers in their assessments.
(4) Adequately blended groups may help to overcome any potential gender and cultural biases.
(5) A significant number of students (about 40%) gave the highest marks possible to their peers across all assessment criteria. This observation suggests that a clear need exists to train the students about the importance and procedures of peer assessment.

Acknowledgement—The authors thank Glyn Mather for proofreading the manuscript. This paper is a revised and expanded form of a conference paper presented at the Society of Petroleum Engineers (SPE) Annual Technical Conference and Exhibition, New Orleans, LA, USA, 4-7 October 2009. The authors thank SPE for the copyright transfer of the original paper to IJEE.

References

1. P. Schenewerk and R. Alman III, *Petroleum Engineering*

Education: Innovation to Meet the Needs of Industry, SPE30655, presented at the SPE ATCE, Dallas, TX, 22–25 October 1995.
2. B. R. Dickson and C. D. Grant, *Engineering Management Skills: The Present and the Future for Technical Graduates*, SPE109108, presented at the Offshore Europe, Aberdeen, U.K. 4–7 September 2007.
3. J. E. Russell, N. J. Broussard and H. A. Tiedemann, *Colloquium on Petroleum Engineering Education: A Summary*, SPE30653, presented at the SPE ATCE, Dallas, TX, 22–25 October 1995.
4. H. Kazemi, W. J. Lee, T. A. Blasingame, R. Alman III, Z. Bassiouni, C. H. Bowman, A. W. Eustes III, D. W. Green, L. R. Heinze, R. N. Horne, J. Judah, M. A. Miller, D. T. Numbere, M. G. Prado and H. Tiedemann, "*The Fifth SPE Colloquium on Petroleum Engineering Education—An Industry Perspective*," SPE64308, presented at the SPE ATCE, Dallas, TX, 1–4 October 2000.
5. S. Bloxham and A. West, Understanding the rules of the game: marking peer assessment as a medium for developing students' conceptions of assessment, *Assessment and Evaluation in Higher Education* **29**(6), 2004, pp. 721–733.
6. M. C. Dyball, A. Reid, P. Ross and H. Schoch, Evaluating assessed group-work in a second-year management accounting subject, *Accounting Education*, **16**(2), 2007, pp. 145–162.
7. R. Dearing, *Higher Education in the Learning Society*, Dearing Report, National Committee of Inquiry into Higher Education, London, U.K. 1997.
8. P. C. Candy, G. Crebert and J. O'Leary, *Developing Lifelong Learners Through Undergraduate Education*, Australian Government Publishing Service, Canberra 1994.
9. A. Walker, British psychology students' perceptions of group-work and peer assessment, *Psychology Learning and Teaching*, **1**(1), 2001, pp. 28–36.
10. B. Spalding, S. Ferguson, P. Garrigan and R. Stewart, How effective is group work in enhancing work-based learning? An evaluation of an education studies course, *Journal of Further and Higher Education*, **23**(1), 1999, pp. 109–115.
11. D. J. Magin, A novel technique for comparing the reliability of multiple peer assessments with that of single teacher assessments of group process work, *Assessment and Evaluation in Higher Education*, **26**(2), 2001, pp. 139–152.
12. V. P. Kommula, J. Uziak and M. T. Oladiran, Peer and self-assessment in engineering students' group work, *World Transactions on Engineering and Technology Education*, **8**(1), 2010, pp. 56–60.
13. K. Willey and A. Gardner, Developing team skills with self- and peer assessment: Are benefits inversely related to team function? *Campus-Wide Information Systems*, **26**(5), 2009, pp. 365–378.
14. I. Van den Berg, W. Admiraal and A. Pilot, Peer assessment in university teaching: Evaluation seven course designs, *Assessment and Evaluation in Higher Education*, **31**(1), 2006, pp. 19–36.
15. S. Cassidy, Developing employability skills: peer assessment in higher education, *Education and Training*, **48**(7), 2006, pp. 508–517.
16. S. Brown and P. Knight, *Assessing Learners in Higher Education*, Kogan Page, London, UK. 1994.
17. R. Murray-Harvey, H. Silins and J. Orrell, *Assessment for Learning: A Guide for Academics*. Flinders University Press, Adelaide, Australia 2003.
18. L. Stefani, Peer, self and tutor assessment: relative abilities, *Studies in Higher Education*, **19**(1), 1994, pp. 69–75.
19. R. Ballantyne, K. Hughes and A. Mylonas, Developing procedures for implementing peer assessment in large classes using an action research process, *Assessment and Evaluation in Higher Education*, **27**(5), 2002, pp. 427–441.
20. D. Boud and N. Falchikov, Aligning assessment with long-term learning, *Assessment and Evaluation in Higher Education*, **31**(4), 2006, pp. 399–413.
21. J. Goldfinch and R. Reaside, Development of a peer assessment technique for obtaining individual marks on a group project, *Assessment and Evaluation in Higher Education*, **15**(3), 1990, pp. 210–231.
22. T. Gatfield, Examining student satisfaction with group projects and peer assessments," *Assessment and Evaluation in Higher Education*, **24**(4), 1999, pp. 365–377.
23. L. K. Y. Li, Some refinements on peer assessment of group projects, *Assessment and Evaluation in Higher Education*, 26(1), 2001, pp. 5–18.
24. M. Lejk and M. Wyvill, Peer assessment of contributions to a group project: a comparison of holistic and category-based approaches, *Assessment and Evaluation in Higher Education*, **26**(1), 2001a, pp. 61–72.
25. C. Brindley and S. Scoffield, Peer assessment in undergraduate programmes, *Teaching in Higher Education*, **3**(1), 1998, pp. 79–89.
26. G. Bushell, Moderation of peer assessment in group projects, *Assessment and Evaluation in Higher Education*, **31**(1), 2006, pp. 91–108.
27. M. Lejk, M. Wyvill and S. Farrow, Group learning in systems analysis and design: a comparison of the performance of streamed and mixed ability groups, *Assessment and Evaluation in Higher Education*, **24**(1), 1999, pp. 5–14.
28. M. Lejk and M. Wyvill, The effect of the inclusion of self-assessment with peer assessment of contributions to a group project: a quantitative study of secret and agreed assessments, *Assessment and Evaluation in Higher Education*, **26**(6), 2001b, pp. 551–561.
29. J. Goldfinch, Further developments in peer assessment of group projects, *Assessment and Evaluation in Higher Education*, 19(1), 1994, p. 29–35.
30. D. Sluijsmans, S. Brand-Gruwel and J. J. G. van Merrienboer, Peer assessment training in teacher education: Effects on performance and perceptions, *Assessment and Evaluation in Higher Education*, 27(5), 2002, pp. 443–454.
31. A. Bilgin and S. Fraser, Empowering students to be the judges of their own performance through peer assessment, *Proceedings of the IASE Satellite Conference on Assessing Student Learning in Statistics*, IASE, Guimarães, Portugal, 19–22 August 2007.
32. T. T Vu and G. Dall'Alba, Students' experience of peer assessment in a professional course, *Assessment and Evaluation in Higher Education*, **32**(5), 2007, pp. 541–556.
33. D. Kember and L. Gow, A challenge to the anecdotal stereotype of the Asian student, *Studies in Higher Education,* **16**(2), 1991, pp. 117–128.
34. P. D. Ladd and R. Ruby, Learning styles and adjustment issues of international students, *Journal of Education for Business*, **74**(6), 1999, pp. 363–7.
35. D. Kember, Misconceptions about the learning approaches, motivation and study practices of Asian students, *Higher Education*, **40**, 2000, pp. 99–121.
36. P. Ling, G. Arger, I. Filonenko, H. Chua and C. Yin, Approaches to study: a comparison of Malaysian and Australian students", in A. Brew and C. Asmar (Eds.) *Higher Education in a Changing World, Proceedings of the 2005 HERDSA Annual Conference in Sydney*, *3–6 July 2005*, Higher Education Research and Development Society of Australasia, Milperra, Australia 2005, pp. 276–286.
37. M. Y. Leung, J. Li, Z. Fang, X. Lu and M. Lu, Learning approaches of construction engineering students: a comparative study between Hong Kong and Mainland China, *Journal for Education in the Built Environment,* **1**(1), 2006, pp. 112–131.

Appendix A

Peer assessment marks before and after Li's normalisation process for a well-performing group

Group Mark= 25

	Student-1					Student-2					Student-3					Student-4					Student-5					Student-6								
Task	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	IER	IWF	FM	Rank
Student-1						4	4	4	3	3	5	5	5	5	4	5	4	4	4	3	4	4	5	5	4	4	4	4	5	4	105.0	1.03	25.8	2
Student-2	3	3	3	3	4						4	2	5	4	4	5	4	4	3	4	4	4	5	4	4	5	4	4	4	4	97.0	0.95	23.9	6
Student-3	4	4	3	3	4	4	4	5	4	5						5	5	4	3	4	5	5	5	4	4	5	5	4	4	4	106.0	1.04	26.1	1
Student-4	3	3	3	4	3	4	4	5	4	5	4	2	5	4	4						4	4	5	4	4	4	4	4	5	5	100.0	0.98	24.6	4
Student-5	4	3	3	3	4	4	4	5	4	5	4	3	5	4	5	5	5	4	4	4						4	5	4	4	5	104.0	1.02	25.6	3
Student-6	3	3	3	4	3	4	4	5	4	5	4	2	5	4	4	5	4	4	3	4	4	4	5	4	4						98.0	0.96	24.1	5

	Student-1	Student-2	Student-3	Student-4	Student-5	Student-6	IER	IWF
Avg effort rating							101.7	
Mean								1.00
Standard deviation								0.04
Rating given to others	83.0	106.0	102.0	103.0	108.0	108.0		
Bias factor	0.82	1.04	1.00	1.01	1.06	1.06		
Normalization factor	1.22	0.96	1.00	0.99	0.94	0.94		

	Student-1					Student-2					Student-3					Student-4					Student-5					Student-6									
Task	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	IER	IWF	FM	Rank	CIM %
Student-1						4	4	4	3	3	5	5	5	5	4	5	4	4	4	3	4	4	5	5	4	4	4	4	5	4	101.4	1.00	24.9	3	3.4
Student-2	4	4	4	4	5						4	2	5	4	4	5	4	4	3	4	4	4	5	4	4	5	4	4	4	4	97.8	0.96	24.1	6	0.8
Student-3	5	5	4	4	5	4	4	5	4	5						5	5	4	3	4	5	5	5	4	4	5	5	4	4	4	106.2	1.04	26.1	1	0.2
Student-4	4	4	4	5	4	4	4	5	4	5	4	2	5	4	4						4	4	5	4	4	4	4	4	5	5	100.1	0.98	24.6	4	0.1
Student-5	5	4	4	4	5	4	4	5	4	5	4	3	5	4	5	5	5	4	4	4						4	5	4	4	5	105.3	1.04	25.9	2	1.2
Student-6	4	4	4	5	4	4	4	5	4	5	4	2	5	4	4	5	4	4	3	4	4	4	5	4	4						99.1	0.98	24.4	5	1.2

	Student-1	Student-2	Student-3	Student-4	Student-5	Student-6	IER
Average effort rating							101.7
Rating given to others	101.7	101.7	101.7	101.7	101.7	101.7	
Bias factor	1.00	1.00	1.00	1.00	1.00	1.00	

a = logistics b = leadership c = group dynamics d = intellectual e = research/writing/editing
IER = individual effort rating IWF = individual weighting factor FM = final mark CIM = change in mark

Appendix B

Peer assessment marks before and after Li's normalisation process for a poorly performing group

Group Mark= 21

Task	Student-1					Student-2					Student-3					Student-4					Student-5					Student-6					IER	IWF	FM	Rank
	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e				
Student-1						2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4	4	4	4	4	2	2	2	2	1	42.5	0.52	11.0	6
Student-2	5	5	5	5	5						4	4	4	4	4	5	5	4	4	5	5	5	5	5	5	4	3	5	4	4	110.5	1.36	28.5	1
Student-3	5	5	5	5	5	3	3	3	3	3						3	2	3	2	3	3	3	3	4	4	4	3	3	3	4	85.5	1.05	22.0	4
Student-4	5	5	5	5	5	4	3	4	4	4	4	4	4	4	4						3	4	5	4	4	4	3	3	3	4	99.5	1.22	25.6	3
Student-5	5	5	5	5	5	2	1	3	1	1	1	1	1	1	1	1	1	1	1	1						1	1	1	1	1	48.0	0.59	12.4	5
Student-6	5	5	5	5	5	4	4	4	3	4	4	4	4	4	4	4	4	4	3	5	5	4	4	4	4						103.0	1.26	26.5	2
Avg effort rating																															81.5			
Mean																																1.00		
Standard deviation																																0.36		
Rating given to others	125.0					67.0					70.0					66.0					93.0					68.0								
Bias factor	1.53					0.82					0.86					0.81					1.14					0.83								
Normalization factor	0.65					1.22					1.16					1.23					0.88					1.20								

Task	Student-1					Student-2					Student-3					Student-4					Student-5					Student-6					IER	IWF	FM	Rank	CIM %
	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e					
Student-1						2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	3	3	3	3	2	2	2	2	1	45.4	0.56	11.7	5	6.9
Student-2	3	3	3	3	3						5	5	5	5	5	6	6	5	5	6	4	4	4	4	4	5	4	6	5	5	111.7	1.37	28.8	1	1.1
Student-3	3	3	3	3	3	4	4	4	4	4						4	2	4	2	4	3	2	3	3	3	5	4	4	4	5	84.6	1.04	21.8	4	1.1
Student-4	3	3	3	3	3	5	4	5	5	5	5	5	5	5	5						3	3	4	4	3	5	4	4	4	5	99.3	1.22	25.6	3	0.2
Student-5	3	3	3	3	3	2	1	4	1	1	1	1	1	1	1	1	1	1	1	1						1	1	1	1	1	44.0	0.54	11.3	6	8.3
Student-6	3	3	3	3	3	5	5	5	4	5	5	5	5	5	5	5	5	5	4	6	4	3	3	4	3						104.0	1.28	26.8	2	1.0
Average effort rating																															81.5				
Rating given to others	81.5					81.5					81.5					81.5					81.5					81.5									
Bias factor	1.00					1.00					1.00					1.00					1.00					1.00									

a = logistics b = leadership c = group dynamics d = intellectual e = research/writing/editing
IER = individual effort rating IWF = individual weighting factor FM = final mark CIM = change in mark

Appendix C

Characteristics of the groups

		Group mark	Intern only	local only	blended	male only	blended	SD	Δ (Bias)
Course 1	1	21			x	x		0.36	0.72
	2	25			x		x	0.04	0.24
	3	27			x		x	0.08	0.19
	4	25	x				x	0.06	0.25
Course 2	5	20	x				x	0.00	0.00
	6	24			x	x		0.02	0.09
	7	26	x				x	0.00	0.00
	8	23			x		x	0.05	0.21
	9	20			x	x		0.14	0.24
	10	26			x	x		0.52	0.55
	11	25			x		x	0.19	0.17
Course 3	12	27	x				x	0.04	0.18
	13	27		x		x		0.02	0.11
	14	26			x		x	0.01	0.12
	15	27		x		x		0.07	0.19
	16	26			x	x		0.21	0.36
Course 4	17	26			x		x	0.01	0.02
	18	25			x	x		0.02	0.09
	19	25	x				x	0.00	0.00
	20	18			x		x	0.08	0.33
	21	24		x		x		0.00	0.00
	22	22			x		x	0.07	0.43

Yildiray Cinar is a senior lecturer in the School of Petroleum Engineering at the U. of New South Wales. Previously, he held various research positions at Stanford U., Clausthal Technical U., and Istanbul Technical U. He holds B.S. and M.S. degrees from Istanbul Technical U. and a Ph.D. degree from Clausthal Technical U., all in petroleum engineering. His teaching and research areas of interest cover core analysis for multiphase flow properties of porous materials, disposal of greenhouse gases in subsurface geological structures, and performance of enhanced oil and gas recovery.

Ayse Bilgin is a senior lecturer in the Department of Statistics at Macquarie University. She holds B.Sc. (Petroleum Eng.) and M.B.A. from Istanbul Technical University, Master of Mathematics (Statistics) by Research and Ph.D. from University of Newcastle, and Postgraduate Diploma in Higher Education (Learning and Teaching) from Macquarie University. She teaches undergraduate and postgraduate students in various topics such as Operations Research, Data Mining and Decision Support Systems. Her research interests include statistics education and applied statistics especially in health sciences.