Author: Weaver, Debbi; Esposto, Alexis
Title: Peer assessment as a method of improving student engagement
Year: 2012
Journal: Assessment and Evaluation in Higher Education
Volume: 37
Issue: 7
Pages: 805-816
URL: http://dx.doi.org/10.1080/02602938.2011.576309

Copyright: Copyright © 2012 Taylor & Francis. This is an Author's Accepted Manuscript of an article published in Assessment and Evaluation in Higher Education, vol 37 no 7 pp 805-816, 2012, available online at: http://www.tandfonline.com/10.1080/02602938.2011.576309

This is the author's version of the work, posted here with the permission of the publisher for your personal use. No further distribution is permitted. You may also be able to access the published version from your library.

The definitive version is available at: http://dx.doi.org/10.1080/02602938.2011.576309
Peer assessment as a method of improving student engagement

Debbi Weaver* and Alexis Esposto
Faculty of Higher Education, Lilydale, Swinburne University of Technology, Victoria, Australia

To encourage increased student attendance and engagement in a third year Economics unit, the curriculum was redesigned to incorporate continuous assessment throughout semester. A component of group project marks were allocated to peer-assessment, in an attempt to address concerns about free-riding colleagues sharing a common mark (Gibbs 2009). This study investigated the consistency of marks awarded to peers within teams, and the acceptance by students of marks awarded by peers.

Students were asked to provide ratings and explanatory comments for each of their group peers. Focus groups were conducted to determine students’ acceptance of this strategy. Eighty student ratings were compared to determine consistency of assessment. Within groups, students who received higher marks from their peers generally awarded marks to their peers across a wider range, whereas students who received lower average grades often awarded the same mark to all team members. These results might indicate that students who were attending class regularly and/or contributing at a higher level were more discriminating in the marks awarded to their peers. Similarly, non-contributors (as identified by their peers) assigned the same or similar grades to each of their peers, possibly due to a lack of knowledge about their peers’ contributions.

Keywords: peer assessment, group work, student engagement, evaluation

* Corresponding author. Email dweaver@swin.edu.au
**Introduction**

A recent review of the literature around group assessment, undertaken by Gibbs (2009) highlighted the issue of fair and equitable assessment

‘Allocating a single group mark to all members of a group rarely leads to appropriate student learning behaviour, frequently leads to freeloading, and so the potential learning benefits of group work are likely to be lost, and in addition students may, quite reasonably, perceive their marks as unfair’ (Gibbs 2009, p1).

Gibbs identified a range of mechanisms for allocating fair marks to students in groups, but also identified other aspects of group work (for example, group size and formation) that impacted on students’ learning outcomes. This introduction discusses the use of peer assessment in student group work, as a means of providing realistic marks to individual students within a group setting.

The use of peer assessment as a tool to monitor and evaluate team performance is well established and has been used extensively across all sectors of the educational spectrum to assess student performance and to encourage, promote and teach student team work skills (Falchikov 1995, Dochy, Segers and Sluijsmans 1999). Definitions of peer assessment abound; however, Topping (1998) defined the practice as

‘an arrangement in which individuals consider the amount, level, value, worth, quality or success of the products or outcomes of learning of peers of similar status’ (Topping 1998, p 250).

Peer assessment has been utilised in different ways by educators, including monitoring the performance of group work, or used as a tool of reflection by students, through to the pragmatic efficiency method of reducing marking workloads for teaching staff (Wheater,
Langan and Dunleavy 2005). The method can take a variety of forms in a range of teaching practices, and include:

- peer nomination, which relates to the practice of identifying the best and worse performers in the group;
- peer rating, which ranks the performance of group members based on a set of performance requirements; and
- peer rankings, which ranks group members from best performers to worse (Pope 2001).

In an exploratory study of peer assessment at a Post-Graduate Diploma in Education in Singapore, Divaharan and Atputhasamy (2002) found that students were positive about group work and the use of peer assessment in tutorial work. The authors found that cooperative learning combined with peer assessment provided a solid basis for the development of several learning skills, including communication skills and support skills, as well as a stronger sense of responsibility towards learning activities. They concluded that the practice of peer-assessment for cooperative learning should be encouraged, particularly with mature students (Divaharan & Atputhasamy, 2002). Similar results were earlier reported by Falchikov (1995) who found that


Dochy, Segers and Sluijsmans (1999), in a review of the literature of 63 studies, also found that a combination of different assessment tasks, including peer-assessment, encouraged students to become more responsible and reflective. Another important finding of this literature review was that peer assessment was effective for summative assessment, “… by
Issues of reliability and validity of peer marking are an ongoing concern raised in much of the literature. As defined by Falchikov and Goldfinch (2000), reliability refers to the agreement between peer ratings, whereas validity refers to the agreement between peer assessment and teacher assessment. In their meta-analysis of 48 studies comparing peer and teacher marks, these authors found that validity was higher when well-understood criteria were used, but that there was no apparent difference in validity between novice (first-year) and advanced students, or across different discipline areas (Falchikov and Goldfinch 2000).

The incorporation of peer assessment into the university curriculum has received much attention in the literature. The views in the literature tend to be favourable, but some question its validity as a learning and assessment tool. Kennedy (2006) contested the claim that peer assessment as a procedure is valid because it introduces a variety of problems and inequities. In his study, he found that the practice was time-consuming, added to the workload of academic staff, and that, contrary to the claim of improved team performance, the practice ended up being counter-productive, causing tensions, distrust and stress amongst group members. He argued that its use raised serious doubts about the validity and reliability of student assessments of peer contributions and concluded that the worthiness of the practice is rarely sound and it should not be recommended (Kennedy, 2006).

Ferguson, Sheader and Grady (2008), studying practical laboratory classes in biology, analysed two major shortcomings of peer assessment. The first is that it can be a time-consuming task when peer assessment is implemented in large classes. The second is that it has been reported as causing unnecessary stress for students (Ferguson, Sheader and Grady...
Their findings showed that combining peer assessment with computer assisted assessment served well as an evaluation tool for an undergraduate practical class consisting of over 400 students. A study on the fairness of peer assessment conducted by Herbert (2006) investigated whether peer assessment discriminated against minorities in mixed teams. Her findings showed no evidence of discrimination either towards females or international students working in teams made up of local and overseas students.

Overall, as Wheater, Langan and Dunleavy (2005) argued, the literature suggests that peer-assessment in higher education is useful both in terms of being an efficient practice as well as having numerous pedagogical benefits to students and educators. However, they cautioned that to ensure its successful and equitable implementation, the practice of peer assessment remains a challenge for academics. In order to succeed it is important to create an environment and practice requiring “… openness in dialogue, good planning, and close monitoring in the early stages” of its implementation (Wheater, Langan and Dunleavy 2005, p. 15).

This paper reports on the introduction of an anonymous peer-assessment component, in an attempt to both gain a more informed picture of the relative contributions of individuals within teams (for more accurate assessment) and to provide feedback to individual group members on their teamwork skills. The objective of this strategy was to consult individual team members for feedback on their colleague’s performance. It was hoped that this would improve the students’ focus on the team work processes, and provide an opportunity for students to take some ownership of their own assessment.

**The context of this study**
Economic Development is a third-year unit for students majoring in Economics as part of the Bachelor of Business at a multi-campus Australian University. At the beginning of this study, the unit was characterised by poor attendance at tutorial classes, low student engagement, and students rarely interacting with each other (or with the teaching staff). In an endeavour to improve this situation and in light of previous successes with a similar unit the previous year (Esposto and Weaver 2008), a process of curriculum redesign was initiated. Major changes were made to the assessment of the unit in 2008, with the aim of attempting to improve attendance, engagement and teamwork skills.

These changes included the introduction of weekly group activities (worth 30% of the semester marks), to be conducted and presented orally during the tutorial sessions. The same groups were also required to work together on a major assignment (worth 40%), with the remaining 30% allocated to an individually-completed essay. Part of the assessment of the group assignment was conducted by peer assessment, within the student teams. The final exam was removed, with all assessment now completed within the 12-week semester.

Student teams were required to make group presentations, in front of both their tutor and other student groups, on a weekly basis on some current issue or a topic drawn from the literature. In order to ensure that teamwork rules of engagement were clearly understood, a workshop on teamwork was organised for the first tutorial to assist students in how to best function in teams, to help develop a clear understanding of what was required in terms of group work input, and also to have a good understanding of negotiation skills in order to address issues related to conflict resolution.

Students worked in groups of 4-6, and groups appeared to get started with their work immediately on group formation in the first week of semester. Groups were encouraged to
self-organise how they would conduct their weekly presentations, so some groups chose to divide the presentation into sections, with different group members presenting each section, whereas other groups nominated a single member to present the work of the entire group. The only stipulation placed on this arrangement by the tutor was that a different group member would present each week.

As can be seen, the introduction of peer assessment is only one of several major modifications to this unit, so any measured change in grades or attendance is likely due to a combination of factors.

**Roles**

The second author of this paper is the Unit convenor, lecturer and sole tutor for this unit, and designed the curriculum changes. The first author is an Academic Developer, and was involved as a consultant during the curriculum redesign phase, and conducted the evaluation of the changes to the unit.

**Implementation of peer assessment**

Our previous experience of strengthening assessed group work had revealed moderate levels of student dissatisfaction with the issue of free-riders (i.e. students who rely heavily on the work and input of others) receiving the same group mark for the project (Esposto and Weaver 2008). In an attempt to deal with this issue and to introduce a fairer marking scheme, students were asked to provide an assessment of their colleague’s contributions to the group assignment. These students had no previous experience of peer assessment, so concerns raised by Ballantyne, Hughes and Mylonas (2002) about the risk of students resenting having to participate in peer assessment too often (Ballantyne, Hughes and Mylonas, 2002) were not a factor. Since this was a novel experience for both the students and the staff, and in light of the
concerns raised by Kennedy (2006), a structured and carefully moderated approach was adopted.

At the end of semester, students were asked to rate (on a simple numeric scale) the contributions of their colleagues in completing the group assignment, with optional additional comments to justify their ratings. This form of peer assessment, encompassing a mixture of quantitative and qualitative judgements, was deliberately used in an attempt to capture the concept of ‘colleague contribution’ in simple terms. Students were asked to make judgements only on the processes involved, rather than the product produced (i.e. their group assignment). Topping (1998) calls this peer assessment of ‘skilled professional behaviours’, in contrast to the assessment of an artefact. Peer assessment was conducted only within groups – that is, intra-group assessment, in contrast to inter-group assessment when students assess the product or presentation of other student groups (Sivan 2000).

Evaluations were submitted privately to the lecturer, so students were not informed of their ratings from peers. This peer-assessment was compulsory, and students who did not submit their evaluations received zero marks for this component.

A portion of the peer-assessment form is shown in Figure 1 (students were asked to provide similar information for each fellow group member). Students could rate a peer’s contribution as a negative impact – i.e. the group would have performed better if that person had not turned up. Most students elected to provide additional comments to support or elaborate on their choice.

It should be acknowledged that although students were verbally asked to assess their peers’ contributions to the processes of group work, this was not clearly specified on the peer-assessment form, and students either interpreted this to mean contributions to the product, or could not distinguish between the process and the product.
Please complete the following colleague assessment and hand back to the tutor.
Name: 
Colleague 1:

<table>
<thead>
<tr>
<th>Liability to group</th>
<th>0 Poor and uninterested work</th>
<th>1 Some work</th>
<th>2 Good Work</th>
<th>3 Very good work</th>
<th>4 Indispensable in completing assignment</th>
</tr>
</thead>
</table>

Comments:

Figure 1: Student peer assessment form

After the group assignments were graded, peer ratings for each group were collated and averaged, and the resulting figure was then applied as a multiplier to the group mark (termed the Input Multiplication Factor, or IMF). The acronym IMF was deliberately chosen to inject a note of humour into this exercise – these students were Economics students familiar with the International Monetary Fund. The rationale of the IMF was to reward the outstanding students, recognising that these students may have achieved a higher mark if allowed to work individually, by awarding these students an IMF > 1.0. Accordingly, students rated as contributing to an average level by their peers were awarded the group mark as their final mark, whereas those who were recognised as contributing more or less than the average received higher or lower marks respectively. The following conversion table was developed for this purpose:

Table 1: Students' Input Multiplier Factor (IMF)

<table>
<thead>
<tr>
<th>Average score (awarded by peers)</th>
<th>IMF score</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.0</td>
<td>1.2</td>
</tr>
<tr>
<td>3.5</td>
<td>1.15</td>
</tr>
<tr>
<td>3.0</td>
<td>1.1</td>
</tr>
<tr>
<td>2.5</td>
<td>1.05</td>
</tr>
<tr>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>1.0</td>
<td>0.8</td>
</tr>
<tr>
<td>0.0</td>
<td>0.6</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.4</td>
</tr>
</tbody>
</table>
We recognise that these figures are somewhat arbitrary, but as this is a pilot project, we thought these figures would give the fairest results for students. As a safeguard however, it was decided not to release these peer ratings to students, but to allow the lecturer to review all ratings, compare these with his own observations, and to make adjustments if he considered any assessments unfair or discriminatory in any way.

The reward system was designed to award students with higher scores when their work was of outstanding quality, not only in terms of the work performed, but also in terms of their input in organising their colleagues’ work, leadership, assistance to colleagues, meeting deadlines, etc. After collation of peer ratings, the lecturer reviewed all summarised scores, checking for any disparities against his own perceptions of individual contributions. If no glaring disparities were obvious, the students’ peer assessment marks were used as part of the formal assessment.

The peer assessment form focussed on evaluating colleagues’ degree of involvement and input in the completion of the group assignment. Students’ marks for the assignment were adjusted (from the mark awarded to the group), depending on the average score attained from student feedback, using the IMF formula presented in Table 1.

**Evaluation project**

To evaluate the effectiveness of this process, a small evaluation project was undertaken across two semesters of the pilot study. Three key evaluation questions were identified:

1. How effective are these changes in improving student attendance?
2. Are students more engaged with the unit?
3. Are there any issues arising from the introduction of the peer-assessment component of group work that need to be addressed?
Methodology:

Data were collected from several sources over the two-year pilot, to attempt to answer these questions:

1. Focus group interviews were conducted with students, after they had completed their peer-assessment ratings, but before final results were made available (1st year only)
2. The unit convenor (and sole tutor) provided data on attendance, and his perceptions of the level of student engagement in class (both years)
3. Student peer-assessment ratings were analysed for consistency and/or discrepancy within student groups (both years)
4. Student grades were analysed, for comparison with previous cohorts (both years)

Towards the end of the first year of the pilot, two focus group interviews were conducted, one for each of the two tutorial groups. The last 30 minutes of a scheduled class was allocated to the interview, and the tutor left the room while the interviews took place. Students were informed that participation was completely voluntary, and that anyone who wished to was welcome to leave (several students did leave in one tutorial group).

Full ethics clearance was obtained from the University’s research ethics committee before commencing any data collection.

Student participation rates

Participation rates for the focus group interviews and for the peer assessment exercise are detailed in Table 2.
Table 2: Student enrolment and participation rates

<table>
<thead>
<tr>
<th></th>
<th>First year of pilot</th>
<th>Second year of pilot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students enrolled</td>
<td>44</td>
<td>44</td>
</tr>
<tr>
<td>Participated in focus group interviews</td>
<td>30 (68 %)</td>
<td>(not conducted)</td>
</tr>
<tr>
<td>Completed peer assessment</td>
<td>43 (98%)</td>
<td>38 (86%)</td>
</tr>
</tbody>
</table>

Results and Discussion

Tutor feedback

The lecturer (and sole tutor) for this unit was asked to provide written comments on his perceptions on whether the changes implemented had achieved the desired outcome, and since this was an initial trial of the implementation of peer assessment, we wished to carefully moderate this process.

The outcomes of the changes implemented were much better than anticipated. As presented in Table 3, student attendance at tutorials in the first semester increased to an average of 84%, compared with an attendance rate of 35% for the previous cohort of students, and this improvement was maintained in the following year. Similarly, student pass rates and grades improved, from a previous average final mark of 60% to 71% (in 2008) and 73% (in 2009).

Of course, these improved attendance and grades result from a range of changes to the unit – and the peer assessment component is only one of those changes.

Table 3: Student grades and attendance and comparison of final marks 2007-2009

<table>
<thead>
<tr>
<th></th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Year 1 of study)</td>
<td>(Year 1 of study)</td>
<td>(Year 2 of study)</td>
</tr>
<tr>
<td></td>
<td>Enrolment</td>
<td>Attendance</td>
<td>Enrolment</td>
</tr>
<tr>
<td>Tutorial 1</td>
<td>31</td>
<td>35 %</td>
<td>19</td>
</tr>
<tr>
<td>Tutorial 2</td>
<td>25</td>
<td>82%</td>
<td>17</td>
</tr>
<tr>
<td>Average final mark</td>
<td>60.3 ± 10.2% (p&lt;0.01 cf 2007 marks)</td>
<td>71.0 ± 20.7% (p&lt;0.0001 cf 2007 marks)</td>
<td>73.3 ± 12.3% (p&lt;0.0001 cf 2007 marks)</td>
</tr>
<tr>
<td>(whole cohort)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The tutor was impressed by the overall standard of the presentation content – students engaged with the topics, identified key issues, and grappled with quite difficult papers. Students appeared to like the subject overall. Few non-performers were observed amongst the groups, and, from his observations, the team work processes appeared to be working well. Competition developed between the groups, leading to debates when groups disagreed on issues, which is a very positive outcome.

**Student feedback**

Questions in the group interviews were focused on the increased emphasis of group work and on the peer assessment component of the assessment. Students reported in the group interviews that they were very happy with the structure of the unit, they enjoyed the group work aspects, and recognised its importance for their employment prospects. There was some concern expressed about free-riding or poorly-performing students creating more work for team members. However, the peer-assessment strategy has meant that students feel these free-riders were not being rewarded for their colleague’s efforts, so no-one appeared overly concerned about this. (This is in contrast to most group projects, where the issue of free-riding students is often a major concern). At the time of these interviews, students had not received their grades for their group assignment, and did not know how these peer reviews would affect their grades, but surprisingly, they did not seem overly concerned with that aspect. It appeared that the mere fact of being consulted about their colleague’s contributions had addressed these concerns.

It must be acknowledged that the positive feedback reported by these students was from those students who attended class and were therefore available for interview. It is likely that students who did not attend class may have had a different view – however the attendance figures were high, so these responses are probably adequately representative.
**Peer assessment**

A total of 80 students, from 16 teams across two student cohorts (Year 1 and Year 2 of the study), submitted feedback on their colleague’s contributions to their group project. This comprised a 98% participation rate in Year 1, and 86% in Year 2. Across both cohorts, the student ratings matched to a high degree the lecturer’s class observations of individual contributions in terms of assignment completion and engagement with colleagues in terms of trying to complete the report, confirming the results of previous studies (Falchikov 1995).

An example of the peer assessment for one group is provided in Table 4. The group participants were awarded a mark of 35 (from a possible 40) for their submitted assignment. This mark was then adjusted for individuals within the group, according to the IMF. As can be seen in this example, one student was recognised by their peers as being essential to the completion of the project, so was accordingly awarded extra marks (beyond the maximum advertised marks for the assignment).

<table>
<thead>
<tr>
<th>Group mark = 35/40</th>
<th>Student 1</th>
<th>Student 2</th>
<th>Student 3</th>
<th>Student 4</th>
<th>Student 5</th>
<th>Student 6</th>
<th>Ave mark from peers</th>
<th>Adjusted mark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 1</td>
<td>X</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.6</td>
<td>42</td>
</tr>
<tr>
<td>Student 2</td>
<td>4.0</td>
<td>X</td>
<td>3.5</td>
<td>3.5</td>
<td>3.0</td>
<td>4.0</td>
<td>3.6</td>
<td>40</td>
</tr>
<tr>
<td>Student 3</td>
<td>2.0</td>
<td>3.0</td>
<td>X</td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
<td>3.2</td>
<td>39</td>
</tr>
<tr>
<td>Student 4</td>
<td>2.0</td>
<td>3.0</td>
<td>4.0</td>
<td>X</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>39</td>
</tr>
<tr>
<td>Student 5</td>
<td>0.0</td>
<td>2.0</td>
<td>1.0</td>
<td>1.0</td>
<td>X</td>
<td>4.0</td>
<td>2.0</td>
<td>35</td>
</tr>
<tr>
<td>Student 6</td>
<td>0.0</td>
<td>2.0</td>
<td>1.0</td>
<td>1.0</td>
<td>3.0</td>
<td>X</td>
<td>1.4</td>
<td>28</td>
</tr>
</tbody>
</table>

Students had been asked to rate their colleagues on the processes involved, rather than the quality of the end product (and the difference between these was explained in class).

However, many students appeared to equate these two parameters, and cited the quality of submitted work (for example, the standard of referencing) as an indication of the quality of
the team work processes. Nevertheless, when examples of this were found, the ratings provided appeared consistent for the level of team work demonstrated.

Students were surprisingly consistent when rating their better-performing colleagues (in the example shown in Table 4, the top-ranked student was rated between 3.0 - 4.0 by all peers), but a wide diversity of ratings was given to those who appear to have been contributing to a lesser degree (in the same group, some students were rated anywhere between 0 and 4.0).

Conversely, the range of marks awarded to peers by the higher-performing students (as determined by their average peer marks) was greater than the range of marks awarded by those who received lower average marks, as seen in Table 5. There was a weak significant correlation (r = 0.25, p<0.05, n=80) between the range of marks awarded to peers, and the average marks received.

This is demonstrated by the wider range of marks awarded by highly-rated student 1 (in Table 4), who awarded marks ranging from 0 to 4 to colleagues. Conversely, student 5 (who was identified as a low-level contributor by peers) gave consistent marks to all other group members. This is further illustrated by the difference in the detail of comments provided by these students about each other:

“Hard to communicate with, handed in his work very late. Half of his work was directly quoted from a website that was not a legitimate information source and he did not reference anything. He really didn’t contribute much because I had to rewrite most of his contribution, and didn’t contribute anything during group discussions. And he wasn’t very motivated in achieving a good mark.” (Student 1 commenting on Student 5’s contribution)

“Worked well in group” (Student 5 commenting on Student 1’s contribution)
Table 5: Comparison of the range of marks awarded against average marks received

<table>
<thead>
<tr>
<th>Range of marks awarded to peers</th>
<th>Average of marks received from peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>2.78 ± 0.68 (n=12)</td>
</tr>
<tr>
<td>1.0</td>
<td>2.77 ± 0.82 (n=28)</td>
</tr>
<tr>
<td>1.5</td>
<td>3.60 (n=1)</td>
</tr>
<tr>
<td>2.0</td>
<td>3.03 ± 0.87 (n=17)</td>
</tr>
<tr>
<td>3.0</td>
<td>3.05 ± 0.84 (n=15)</td>
</tr>
<tr>
<td>4.0</td>
<td>3.14 ± 0.84 (n=5)</td>
</tr>
<tr>
<td>5.0</td>
<td>4.00 ± 0.00 (n=2)</td>
</tr>
</tbody>
</table>

Note: Students could award marks from -1 to +4, giving a possible range of marks of 5.

Hence, higher performing students demonstrated greater discrimination in their marking, when compared to their lower-rated colleagues, and gave marks over a greater range of values. These results can indicate that either or both of two things are happening:

1. Higher-performing students more likely to give reliable / credible assessment of their peers’ contributions, and/or
2. Non-performing students (who may also be serial non-attendees) are ignorant of the relative contributions of their peers, so take the ‘safe’ option, and award similar marks to all colleagues.

These results are similar to those reported by Falchikov and Boud (1989), and also support to the results reported by Lew Alwis and Schmidt (2009), looking at reliability of self-assessment ratings:

‘…students judged as more academically competent were able to self-assess with higher accuracy than their less competent peers’ (Lew, Alwis and Schmidt 2009, p. 151).

When our student peer-assessment grades were averaged for each individual, the resultant marks tallied closely with the tutor assessment of individuals’ performance, for both years of the study, and the tutor determined that no further adjustments to marks were required.

16
Providing additional comments on the peer assessment form was optional, however most students took this opportunity to provide detailed explanations or justifications of their ratings.

Conclusions

The implementation of a peer-assessment component into the assessment of group assignments has proved popular with students. This strategy was initially implemented as a result of consultation with students in a previous study (Esposto and Weaver 2008). In that study, students’ only concerns with an increased focus on group work was about equal marks being awarded to non-contributors. The strategy is also in line with institutional pressure to provide feedback to individual students on their progress towards attaining key generic skills (including teamwork skills). In the current study, students were very positive about being consulted on the relative performance of peers, and appeared unconcerned about how their assessment of their peers would be used, contrasting the results of Kennedy (2006). This lack of concern is most likely linked to the fact that the peer assessment was not revealed to colleagues – most students stated that they would be less comfortable about providing similar feedback directly to their peers.

In this study, we made a deliberate decision not to show students the feedback from their peers – primarily since this was a pilot study to determine acceptance and reliability first, and partly to assuage tutor concerns about the processes involved. We intend to provide collated peer comments back to students, as recommended by Falchikov (1995), and in line with our institution’s objectives of providing ongoing feedback to students on their development of key generic skills (with team work skills being explicitly included here.) To be successful, this would require some training and guidance for students in how to constructively word their feedback to their colleagues:
“…it can be argued that learning how to make criticism in a diplomatic manner is a useful skill in itself” (Falchikov 1995, p. 184)

This pilot study of incorporating peer assessment has proved more successful than was hoped, and has given us confidence to proceed with the next stage of providing peer feedback to individuals on their teamwork performance. We believe this should be anonymous and averaged across group members, and will need careful moderation from the tutor before being released to students. However, we believe that feedback from peers, who have greater insight into the inner workings of the groups, will provide a stronger student–centred focus and thus carry greater weight than feedback from a tutor, necessarily situated outside the group.

While it is difficult to gain accurate evidence of the level of student engagement with a unit of study, the feedback from both the student interview groups (conducted in the first year of the pilot study) and from the tutor were positive, and in accord with each other, and hence are likely to give a reasonable indication of the level of student interest.

Based on our experience, peer assessment should be considered as an integral component of group work, but does need moderation to ensure only students who are active participants in the group activities (and therefore more likely to be informed about relative contributions of peers) are eligible to provide assessment of their colleagues, to minimise the impact on grades from uninformed non-attendees. How this might be effectively managed is a topic currently being considered – it is possible to exclude the peer ratings of students identified by their colleagues as non-contributors, but this is a very labour-intensive task. Possibilities using peer-assessment software are current being investigated.

Acknowledgements
The authors acknowledge with gratitude the friendly co-operation and helpful feedback provided by students undertaking this subject during 2008 and 2009.

Notes on contributors

Debbi Weaver is an Academic Developer in Swinburne Professional Learning, at Swinburne University of Technology, Melbourne Australia. She works alongside faculty members on a wide range of teaching innovations and evaluations.

Alexis Esposto is a senior lecturer in Economics, in the Faculty of Higher Education, Lilydale, at Swinburne University of Technology, Melbourne Australia. Alexis' main interests are in measures of human capital and the concept of skill, employment and the ageing workforce, labour market change, inequalities in the workplace, and the global knowledge economy.

References


