
Further information on conference website:
http://www.ics.heacademy.ac.uk/events/11th-annual-conf/index.php

Publisher’s copyright statement:
This paper is © the Higher Education Academy Subject Centre for Information and Computer Sciences, 2010, and is reproduced with their permission.

For more information and to view the definitive version of the proceedings of the conference, see:
http://www.ics.heacademy.ac.uk/events/11th-annual-conf/index.php

Always use the definitive version when citing.

Use Policy:
The full-text may be used and/or reproduced and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not for profit purposes provided that:

- A full bibliographic reference is made to the original source
- A link is made to the metadata record in Newcastle E-prints
- The full text is not changed in any way.

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.
ABSTRACT
A student’s learning achievements are still interpreted, largely, in terms of their quantitative scores on assessment tasks completed during or at the end of a module of study. Associated with this interpretation is the somewhat tacit assumption that, if a student scores well on assessments, they must be highly engaged with the learning methods and materials associated with that module and their level of engagement has contributed to their success. In this paper, we reflect on peer assessment methods used in group work and try to determine if and how these can impact on student engagement and academic achievement. We review and analyse the quantitative and qualitative results from two undergraduate modules that use different peer assessment methods. We try to determine how the use of peer assessment in each of the modules has impacted on students’ engagement, their learning outcomes and their overall academic achievement. We then discuss how greater engagement could be fostered using a wider range of peer and self-assessment methods and we also give examples of how these methods can be incorporated into current assessment frameworks to facilitate greater student learning and a more accurate measurement of student achievement.

Keywords
Peer assessment, student engagement, academic achievement

1. INTRODUCTION
Team work assignments have become the norm in most HEI undergraduate programs for the sound pedagogical reason that university education is not just about developing what people know and understand in isolation but about learning from and with others, for the benefit of society. Teamworking helps students “shift away from simple academic achievement to much broader goals – preparing them for their working lives” (Leik et al, 1996). However, assessing individual performance in teamwork can be complicated because we need to measure the effectiveness of the team process as well as the product/s delivered at the end of the learning activity. Accreditation and certification requirements ultimately mean we have to generate a summative mark that fairly represents the effort each student has put into the team activity and the level of success they have achieved based on the expected standard for their level of study. However, converting a student’s contribution on a group task into a numeric grade “is a complicated and problematic task”, (Leik et al, 1996). At Newcastle we have often used peer assessment as one way to understand the team process more accurately. We view it as a good mechanism for getting a more rounded picture of how teams have interacted both formally and informally when trying to get their work done. With any assessment, the mark given is the final interpretation of learning achievement but often associated with this mark is the assumption, somewhat tacit, that if a student has performed well and achieved a high personal score, they have been highly engaged in the teamwork process and this level of engagement has contributed to their success. But is this true? We wanted to know if the peer assessment methods we use were accurate measures of a student’s contribution to teamwork and their engagement in the activity. We also wanted to find out if the methods themselves had impacted on the students’ engagement in any way. In this paper we review and analyse the quantitative and qualitative results of two different methods of peer assessment used in our undergraduate teamwork activities. We try to determine the impact of these methods on students’ engagement with the activities and their overall academic achievement. We then outline some alternative methods of peer and self-assessment that could be used to increase engagement and give examples of how these methods can be incorporated into current assessment frameworks to facilitate greater student learning and a more accurate measure of student achievement.
2. **Peer Assessment at Newcastle**

2.1 **Method 1: Allocating Percentages**

The Software Engineering Team project at Newcastle is a year-long project activity undertaken by our level 2 students. Students are formed into teams and must create a large piece of software and documentation by the end of the academic year. Student teams are expected to work almost independently but they are given support when needed. For peer assessment in this project we ask students to divide 100% between their team members based on their perception of how much effort and contribution to the project processes and deliverables each team member has made. Students complete the peer assessment twice during the project, once, quite early on in the project – approximately halfway through the first semester and again at the end of the second semester when all the major deliverables have been submitted and demonstrated. For both instances students are asked to distribute 100% between team members, including themselves. An example percentage allocation for a 5 person team is illustrated in Figure 1:

![Figure 1: An example of percentage allocation](image)

In this example, the total column actually adds up to more than 100% for the whole team but this fact is largely deemed insignificant, it is hard to be precise but we round up above 0.5. The marks for all team assignments are then multiplied by these two peer assessment marks to get the student's overall individual mark for the team deliverables e.g. if a design document received a mark of 77%, Student 1 would receive 22.6% of the group mark of 77% for the document = 17.04. We use both individual and team assignments during the module but also **monitoring** and **observation** of teams by a member of staff. The member of staff allocated as a monitor attends each team's weekly formal meeting and assesses student's effectiveness during the project, as a team and as an individual, based on their performance in meetings throughout the year. It is the combination of marks from individual assignments, team assignments, monitor observations and peer assessments that are used to calculate an individual's module mark at the end of the year.

2.2 **Method 2: “Show Me the Money!”**

The second project in which students peer assess is a video filming project that students undertake in the Information Handling module. Their task is create a 5 minute video relating to their chosen theme e.g. Student Life. The aim of the project is to illustrate the legal issues that arise e.g. copyright of images, filming in public places but also the technical issues of film editing and of course the important aspect of working in a team. Students are placed into small teams of 5-6 students and asked to produce a series of assignments relating to their final video submission. These assignments include writing storyboards and shooting scripts. As part of the module students are asked to review their team process via peer assessment. They are asked to distribute an imaginary £1000 bonus between group members according to their perception of each person's effort and contribution to the group goals. Each student allocates a proportion of the £1000 to their team mates again, including themselves. The tutor then finds the maximum amount allocated to one student and that student gets all the mark given e.g. 20 out of 20. Everyone else gets the marks times their total divided by the maximum amount. So, for example, if a maximum mark is 19/20 for a piece of team coursework the allocations for a team of 5 students would be as follows in Figure 2. In this example students B, C and E would all get the full 19/20 and students D and A would get 17.9/20 each.
Figure 2: Allocation of money in Information Handling

3. IMPACT ON ENGAGEMENT & SUCCESS

Carini et al, in an analysis of student surveys from 14 institutions and 1352 students defined some student engagement indicators as positive views on

- the level of academic challenge in a programme,
- the level of student-staff interaction,
- a supportive climate
- the level of active and collaborative learning.

They found that whilst student engagement is linked positively to learning outcomes such as critical thinking and grades the relationship between engagement and academic performance was not as robust as they had hoped. Student engagement constitutes a constellation of institutional processes that may add value to student learning, (Carini et al, 2006). At Newcastle, both the Software Engineering and Information Handling modules have mainly received positive student feedback. Students enjoy doing something different, they like working in teams and they are eager to take on the challenge of creating a video or a large piece of software. Student performance is also generally what is expected within a level 2 class. There are however, quite a few disagreements in teams when it comes to allocating marks for peer assessment and students often complain about this process to module leaders. We want to change peer assessment so that students have greater engagement with their team projects and to ensure the process itself helps students with their learning.

We have been using the percentage sharing method in the Software Engineering Team Project for over 5 years. We compared a small sample of module marks from 12 teams during one academic year (2008/09). There were 106 students in the class. We examined the interim and final peer percentages for the module to see if there was any indication that peer percentage was a reflection of the student’s overall mark for the module. The average team size was 9. The lowest interim percentage mark give was 1.55% whereas the lowest final percentage was 0. The highest percentage given to a team member at the interim was 19% and the highest mark for the final set of peer assessments was 24. There was a lot of variation between teams in the second semester, compared to the first. We found that students were more inclined to give each other equal marks in the first half of the project when only one major deliverable is submitted i.e. the requirements document. The second semester is when the project enters the major implementation period. This is often a tense time for the students, especially if they have not stuck to their original plan or have poor team cohesion or morale. We found that there was a wider range of marks used by each team for all second semester percentages and that peer assessments strongly reflected students’ final mark in the majority of cases. Also 8 of the 10 best performing students in the class received a poorer rating from their peers in the second semester, with 4.87 being the biggest percentage fall. This result has led us to believe our design and schedule for the project may have caused unnecessary stress and contributed negatively to performance and peer assessment outcomes during the second semester.

We have only been using the peer assessment in the video project for one year and therefore believe it is too early to draw any conclusions on the quantitative results. One interesting observation we have for this class is how the use of money rather than marks in the peer assessment has affected students’ approach to peer assessment. Anecdotally, we have found that students are much more focused on the assessment task when it comes to allocating money and less focused on their overall module mark and the outcome of the project in terms of their final grade. Students also seem to be more objective in their allocations. We have observed that money is less emotive for them than marks and they are inclined to be honest and forthcoming about their own contribution when it comes to a discussion on how the £1000 should be distributed. We think that this is because the money is seen as unconnected with their final degree classification and is viewed as more of a bonus mechanism that graduates could encounter in their working lives on a real project. This is in complete contrast to how peer percentages are viewed in the Software Engineering project where discussions often get

<table>
<thead>
<tr>
<th>Student A</th>
<th>(£200) (£200) (142.86) (£200) (£200) 942.86 17.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student B</td>
<td>(£200) (£200) (£200) (£200) (£200) 1000 19.0</td>
</tr>
<tr>
<td>Student C</td>
<td>(£200) (£200) (£200) (£200) (£200) 1000 19.0</td>
</tr>
<tr>
<td>Student D</td>
<td>(142.86) (£200) (£200) (£200) (£200) 942.86 17.9</td>
</tr>
<tr>
<td>Student E</td>
<td>(£200) (£200) (£200) (£200) (£200) 1000 19.0</td>
</tr>
<tr>
<td>Student A’s mark: 19 * 942.86/1000 = 17.9</td>
<td></td>
</tr>
</tbody>
</table>
very heated when it comes to allocation of marks. There is a reliance on students to be honest for this peer assessment and although we think students are honest, some safety measures need to be taken to maintain academic integrity and quality standards. One such safety measure could be the use of contribution matrices, (Devlin et al, 2008). A contribution matrix is a grid where each task the team completes is broken down into sub-sections and students must note down all the sub-tasks they contributed to133(133,11),(992,991). The matrix is then agreed by all team members and submitted along with the assignment/deliverables. We are going to observe the allocation of money in the Information Handling module next time round to see if our observations are repeated. We believe the move away from allocating marks on this peer assessment task has had a positive effect.

3.1 Feedback from students

In both projects, peer assessment is used as a summative measure of contribution towards the end of the learning activities. However, in the case of Software Engineering Team project, peer assessment is also used as a formative process in the first semester. It is the summative peer assessment that seems to cause the most friction between teams. Feedback from both modules on the peer assessment exercises has been quite negative as can be seen in the Figure 3:

```
"I was not happy with my original peer percentage for the second semester which was 14.4 mostly because I felt like I had done more work than some others. After listening to my objections, my mark was increased by the team. However, this meant that two other students lost some marks as a result. Obviously they were not happy about this. Eventually we all (reluctantly) agreed on a mark that was fair. Personally I don't think this is a good system as it can be abused easily, for example two members could unfairly rate each other."
```

Figure 3: Student Comment

This student who commented in Figure 3 took part in the Software Engineering project. We specify that students should conduct the peer assessment exercises during formal meetings when their monitor will be present. This is one way we try to ensure that members are not forced to agree to a mark that is unfair but also, because we try to avoid the situation of collusion that this student mentions. Unfortunately, we cannot be fully sure that collusion has been avoided. There is also some evidence of students performing differently in observed and unobserved team meetings. Other student comments have made us think about the assignment tasks that we set for the students. We ask students to choose team roles based on their skill strengths and weaknesses, (Devlin et al, 2007), but the choice of roles and tasks means that some students may not get a fair allocation in the first set of peer assessment percentages, through no fault of their own. Evidence of this impact can be seen in the statement in Figure 4 where the scheduling of tasks and assessment deliverables has had an impact on the first set of marks some students receive. This may account for some of the differences between ratings in the first and second semester for the high-scoring students mentioned earlier.

```
"At the moment much of the work is technical and it is clear that certain members of the group have stronger ability in this area. We believe that as the project progresses each team member will have more of an opportunity to stand out and achieve a higher personal percentage."
```

Figure 4: Comment on Module Design

Some teams also noted that peer assessment had a demoralizing effect on the team and affected relationships, which was not really optimal for the continuation of work during the second semester. The reason we introduced peer assessment in the semester of these projects was for the purposes of giving formative feedback to students i.e. to help further learning and to allow them to modify their behaviour to improve their learning outcomes during the project and ultimately their final mark at the end (Freeman et al, 2006). The use of peer assessment at the end of each of the projects was, in part, to help us estimate an individual’s contribution for summative purpose, but mainly to help students reflect on their performance and what they have achieved and learned. However our students’ experiences of peer assessment are somewhat negative and we need to improve this. It may be that project scheduling, the methods we use for peer assessment, (especially the one for allocation of percentages) and the methods that we use for monitoring performance during meetings are not working as well as we had hoped in combating collusion but also in terms of engaging students with the process and reassuring them about fairness. More importantly the
methods are not helping students to focus on their learning or to assess their skills development throughout the project.

4. GREATER ENGAGEMENT VIA PEER ASSESSMENT

According to Falchikov, when it comes to peer assessment, students need guidance, (Falchikov, 1988). One of the problems with our methods is that we do not give our students much guidance on peer assessment. We do give them tips about making fair decisions e.g. we ask students to focus more on tasks, attendance, performance rather than personality traits and behaviour. However, this is not enough, we admit that. Students are also very mark oriented and any assessment should be closely tied to feedback and aimed at motivating students (Falchikov, 1988). We do not give them any guidance or practice on peer assessment before they have to do it for real. We ask students to judge a person’s performance but do not give them any real criteria to do so. Students are therefore unsure of the reasons for peer assessment and find it difficult to do. Some ways to increase students’ engagement with the peer assessment process is to illustrate the real value placed on evaluating performance as a skill needed in the real world. We could make the process more realistic in terms of an appraisal process students will experience in their future careers. Goldfinch and Conway both suggest the use of matrices outlining expected task performance and behaviour but both these approaches still ask students to assign values and marks to their peers and Goldfinch’s approach uses negative marking, which might make peer assessment more emotive than it is already, (Goldfinch, 1994, Conway et al, 1993). Alternatively, Smith and Smarkusky’s approach to peer assessment is similar in that they use matrices however, their competency matrices focus on skills and competencies and ask students to rate themselves using common degree classifications rather than actual marks, (Smith and Smarkusky, 2005). This is almost a variation on our money distribution approach, (see Figure 6). These matrices or an adaption of them would help to capture learning and give students rich feedback on their abilities and skill levels more than our current approaches.

<table>
<thead>
<tr>
<th>CLASS RANK</th>
<th>1ST</th>
<th>2ND</th>
<th>3RD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process</td>
<td>Steps required to complete project</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task Performance</td>
<td>Exhibits on tasks behaviour consistently</td>
<td>Supports others in completing tasks</td>
<td>Motivates others to independently stay on task</td>
</tr>
<tr>
<td>Leadership skills</td>
<td>Learns about leadership skills</td>
<td>Rehearses leadership skills</td>
<td>Exercises leadership skills</td>
</tr>
</tbody>
</table>

**Figure 5: Smith and Smarkusky’s Competency Matrix**

A competency matrix captures team knowledge and skills in various categories (process, communication, interaction, contribution and responsibility). These matrices then allow an assessor to assign a numerical range of proficiency in each specified competency – individuals are evaluated by selecting a class rank to indicate the baseline competencies expected of the individual. In this work, peers assess whether an individual has met the expectations, exceeded the expectation by various amounts or requires improvement (varying amounts of improvement can be denoted). Using a similar approach would enable students to focus on behaviour during the project and the allocation of marks would be left to the tutor. Students would have to evaluate themselves and the performance of others in their team but not assign grades. This would support formative learning better and remove the emotive aspect that assigning marks often evokes. Students could be provided with examples of how to assess in a workshop before the project commences. We could show them how to approach a peer review and how to fill in the matrices for themselves and others. The use of competency matrices would also introduce an important element of self-assessment, early on in the projects and give students the language to assess their skills and performance throughout the year. Staff could use the same matrices and criteria to assess students as the students themselves and then compare matrices in order to assign grades based on a scale defining if they have met expectations, exceeded expectations or if they need to work on certain areas. This would give a sense of continuity, a transparency in assessment and a shared language for staff and students to discuss performance. This method would also give students some experience of ‘social comparison’. Students very rarely get to view the results and feedback given to others and therefore find it hard to compare their performance to that of others in their class in a formal or managed way. The method would also avoid the collusion problems we have experienced and help students to focus on the learning or ‘apprenticeship’ aspects of a project. We could use the matrices periodically to get students to evaluate how they think they are performing along a set of pre-defined competency areas. For software engineering and information handling students we could use competency matrices based on subject-specific competencies identified by Turley and Beiman, for example on “using knowledge, researching, investigating and problem-solving, communication and developing solutions”, (Turley and Bieman, 1995), but the work
outlined here could be adapted quite readily for other disciplines and not just Software Engineering or Information Handling. The matrices would be used in conjunction with the traditional assessment of technical work products and tangible deliverables such as code and documentation.

5. CONCLUSION
All team project experiences are designed to give students a realistic experience of working on a large piece of software and an insight into what it is like to work on real problems within development teams. The alternative methods of peer assessment we outline here offer an opportunity for students to receive higher quality feedback on their progress and development and to determine how to further develop their skills without focusing on grades which may help with their engagement in peer assessment and in their overall project tasks. The current matrices that we have proposed focus on skills and knowledge needed in our modules but it is feasible to identify generic skills and competencies required in team working scenarios for other disciplines. Rarely does a professional person work in isolation these days and generic skills such as communication, leadership, negotiation and problem-solving are required of most professionals in the modern workplace. These generic team working skills could be used in competency matrices that focus on skills rather than scores and substituted for the peer assessment methods we have outlined here. We believe that the focus on skills rather than marks might help students engage more with their learning and with peer assessment exercises which should have an impact on their learning outcomes and achievements.

6. REFERENCES
Falchikov, N., (1988), Self and peer assessment of a group project designed to promote the skills of capability, Programmed Learning and Educational Technology, 25(4), November 1988, pages 327 – 339