Paper

Examining Peer Assessment Methods Amongst Undergraduates: A Student Review

Heather Collis[1], Calum Ridyard[1], Michael Grove[2], Rosemary Dyson[2] and Chris Good[2]
[1] School of Mathematics, University of Birmingham (Student)
[2] School of Mathematics, University of Birmingham (Staff)

Abstract
This report examines the current methods of peer assessment used within the School of Mathematics at the University of Birmingham. Here the focus is upon the use of peer assessment when deciding how to allocate marks between different members of a group to reflect their individual contributions to group work rather than the peer marking of individual assignments and tasks. It uses both quantitative and qualitative responses from students who have experienced one or both of the methods currently used within the School. Further, it goes on to discuss and compare alternatives to the current methods used and proposes possible improvements to be used by the School.

Introduction

In the modern world, undergraduates are now expected to graduate not only with great technical knowledge and understanding of their subject, but also with skills that make them more employable. These skills include organisation, decision-making, presentation, and teamwork (McCorkle *et al.* 1999). Group work – an assessed task where a group of students work together over multiple timetabled sessions, that also requires work to be completed outside of the normal class meeting – has long been thought of as an effective method of developing such skills (Jacques, 1984).

Assessment of group work, however, raises an issue when it comes to distributing marks fairly to individual group members (Conway *et al.* 1993; Cheng & Warren, 1997). Possible choices for the module leader include: (i) simply marking the projects and awarding the same grade to each member of the group; (ii) assigning members of the group different tasks within the project and grading them individually on their own part of the whole; or (iii) allowing groups to explain the division of workload and somehow incorporating this into each individual's mark; this is sometimes called intra-peer assessment (Race, 2001). Cheng and Warren (1997) make the case that awarding the same grade to group members can seem unfair (Cheng & Warren, 1997). Similarly, Goldfinch and Raeside (1990) point out that having students work on individual portions of a project *'disrupts the spirit of group work'* leading to reduced development of personal and interpersonal skills, effectively making the collaborative aspect of the task pointless (Goldfinch & Raeside, 1990). As such, it would appear that the third option would be the one to pursue – handing part of the assessment process over to the students and finding some way to incorporate this with the module leader's assessment of the work – so called 'peer assessment'. A range of peer assessment options are available: students can assess themselves and each other on technical ability and the ability to work in a group, and groups can be assessed on productivity and collaboration (Webb, 1997).

Conway *et al.* (1993) describe in more detail the methods of peer assessment shown in Figure 1. *'Pool of marks to be divided by negotiation'* involves a project being awarded a mark and then the students deciding how much of that mark to give to each person (Conway *et al.* 1993). This is fair, as students are the ones that understand who has contributed most to the project. However, the project marks awarded to each person are not comparable, because some groups may have ranked different elements of the project work as more important compared to other groups. This can cause consistency issues with marks throughout the module. Therefore clear criteria are needed for the allocation of marks.

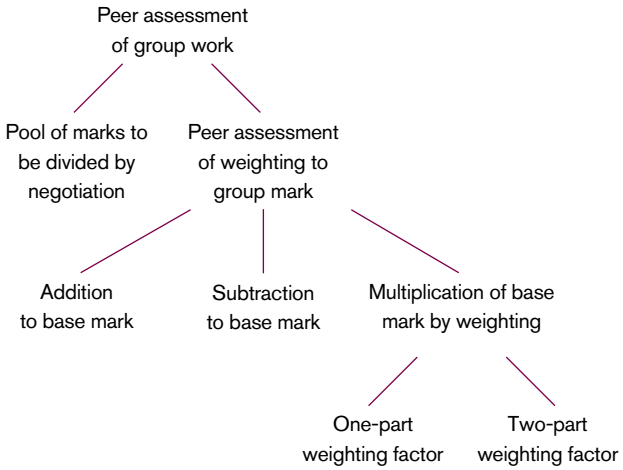

Figure 1: Methods for assessing individual contributions to group activities (Conway *et al.* 1993).

'Addition/Subtraction to base mark' is where each student is awarded the same mark for their project and they may gain or lose marks depending on their contribution to the group (Conway *et al.* 1993). The amount added/subtracted can be decided either by the module leader or the group themselves. This method means students' marks are comparable, as the marks awarded for the project are based on the same guidance. Often, this method doesn't create a significant difference in students' marks, which may cause issues with 'free-riders' not being penalised strongly enough.

'Weighting factors' allocates each student with a proportion of the mark awarded to the group project (Conway *et al.* 1993). This method differs from the *'Pool of marks to be divided by negotiation'* as the weighting factor is calculated by students critiquing each other against a rubric. As students are responsible for awarding the marks used to calculate the weighting factor, it makes them more accountable for their actions throughout the projects. This also means that students' marks are comparable across the whole module, as well as marks being varied enough that 'free-riders' will no longer be successful in the module.

Methods

This report samples first and second year single honours Mathematics students at the University of Birmingham. All students undertook modules which contained, in some part, tasks requiring them to work together to solve problems. At the end of the tasks, members of each group evaluated their own and each other's work within the group. These evaluations were then given to the module leader to be considered in the marking process, and students were all made aware that this would be the case at the beginning of the module. A sample of 55 students were questioned at

the end of the academic year (2014/15) with specific references to their confidence in peer assessment before and after completing the module and the usefulness of peer assessment within the module.

More specifically, the first year students took a 10 credit Mathematical Modelling & Problem Solving module (MMPS) in their first semester. It required them to work in randomly assigned groups of 4 or 5 and complete 4 projects over an 8 week period. At the end of each project, the individual group members all filled in an online diary on the University's Virtual Learning Environment (VLE), in which they discussed the process and division of workload for that project, and evaluated their performance as an individual and as a group. There was also a small numerical component to the diaries where students provided outline allocations of the contributions made by both themselves and their peers to the task. Appendix 1 shows the prompts given by the module leader to students in order for them to complete their diaries. The module leader then allocated marks based on their performance, which, in total, were equivalent to 20% of the module mark, but independent of the project grades. This method is essentially *'Addition to base mark',* as shown in Figure 1. At the end of the academic year, a sample of 26 first-year students were questioned.

The second year students took a 10 credit MMPS module similar to the one in the first semester of their first year – the primary difference being the lack of a numerical component to the diaries. They also took a 10 credit Mathematics in Industry module in the second semester of their second year: 65% of the marks for this module came from two group projects, both of which included peer assessment to allocate the marks to each individual in the group. The larger project, accounting for 45% of the overall module mark, involved students being randomly assigned to groups of 4 or 5 students in which they completed a substantial industry-style research project; the smaller project (20%) allowed students to choose a group of 2 to 5 members in which they worked to produce a resource to communicate mathematics to a wider audience. At the end of each project, each group was required to fill in a form (see Appendix 2) allocating a numerical value to each team member for a list of criteria (similar to that described by Conway *et al.* 1993). This gave every member a percentage corresponding to the work they had put into the project. This was then used to give them the corresponding percentage of the marks for that project. Again, this can be seen in Figure 1 as a *'One-part weighting factor'.* At the end of the academic year, a sample of 29 second-year students were questioned.

Both sample groups were asked if they had experienced peer assessment before coming to university and how this experience (or lack of) affected their view of peer assessment before they started. They were asked to comment on the appropriateness of the types of peer assessment they experienced, and to what extent they understood how the peer assessment would be considered in their grading. The second year students who had experienced more than one method were also asked to compare the two. Finally, there was an opportunity on the form to give any suggestions for improvement of the peer assessment methods.

Review

From Table 1, it can be seen that many students had neutral and slightly negative opinions of their own ability to peer assess, whether they had prior experience or not. The high respondent percentage (33 out of 55 students) for this region comes primarily from two strata – first year students with prior experience (contributing roughly 20% of the weighting) and second years without prior experience (contributing roughly 50% of the weighting). Both groups may have been underestimating their ability, but there could be other reasons. Students without prior experience would naturally feel less confident. However, the reason behind the first year group's response is less obvious, especially as their less experienced counterparts felt more confident (with two of them marking themselves as very confident) – perhaps some of them had negative experiences with peer assessment at school. Harris and Brown (2013) note that no form of peer (or self) assessment at school is anonymous and this can create *'threats to psychological safety and inter-personal relationships'* as peer pressure is more prevalent in secondary schools. These issues may affect students' perceptions of peer assessment as they progress into higher education and, as a result, their confidence in their ability to peer assess may decrease.

Overall, the students were fairly evenly split when asked about how appropriate the peer assessment was for their course (the results can be seen in Table 2). First year students tended to lean towards a negative viewpoint, whereas second year students had a more favourable view. Focusing on the responses of the second year students, it is worth noting that students didn't necessarily vote 'yes' for one style of peer assessment and 'no' for the other (although the results in Table 2 do imply that). There were 10 students that voted 'yes' and 'yes' and 8 students that voted no and no. This suggests that there may not be one method of peer assessment flexible enough to be used universally for all assessments and that a combination of methods may be required to obtain the best results. For example, peer assessment comprised of a reflective section and a numerical section, similar to the method used by Goldfinch (1994), would provide the most flexibility.

When asked to explain the reasoning behind their answers, students gave varying responses. This is expected given how evenly split their opinions are. Many students felt the anonymity of the first year diaries allowed for fairer marks as students could openly discuss each member's contribution to the group. Second year students took issue with the group discussion required for the second year module and would have preferred an anonymous method, suggesting the use of the university's VLE:

'*...the scoring [for Mathematics in Industry] should be anonymous rather than in front of the group. Maybe this could be done on Canvas [the university's VLE].*'
Second year student

	First Years		Second Years		Total
	Prior experience	No prior experience	Prior experience	No prior experience	
Very confident	0	2	1	0	3
Slightly confident	3	2	2	4	11
Neutral	3	2	2	8	15
Slightly unconfident	6	3	1	8	18
Very unconfident	0	5	0	3	8

Table 1: Student perceptions of their ability to peer assess prior to taking the first year module.

	First Years	Second Years	
	Diary-Style	Diary-Style	Numerical-Style
Yes	12	16	15
No	14	13	14

Table 2: Responses to the question: 'Was the style of peer assessment appropriate for the module?'

The method used by Brutus and Donia (2010) of a centralised online peer assessment format is an example of submitting peer assessment *via* a VLE. It allows for staff to easily track students' progress in developing group-related skills and can be tailored so that students can provide their feedback anonymously (Brutus & Donia, 2010). In contrast, a number of students commented that they appreciated the opportunity to justify the numerical values they awarded in a face-to-face environment, but would like some more structure:

'I know other subjects allocate a time for the group to meet with their project leader who leads the group discussion...'
Second year student

The second year students were also asked which method of peer assessment they preferred (see Table 3). The majority favoured the numerical method used in the second year module; very few preferred the diary method, but many had no preference. This shows that there may not be one style of peer assessment that will fit all needs, as well as highlighting the importance of incorporating a written element, alongside a numerical element, into the peer assessment process.

Diary	4
Numerical	15
No preference	10

Table 3: Responses to the question 'Which method of peer assessment did you prefer?' (second years only).

However, some students mentioned that the written element of the process was time consuming:

'With a large project to do, it just seems a hassle to write a diary.'
Second year student

A method to combat the above issue, used by De Wever *et al.* (2011), would be to include a rubric for students to assess against. The rubric could provide detailed written statements that students then have to score their peers against. This avoids the issues of students having to write some text, but still allows for a more detailed review of their peers' contributions than a simple numerical method alone (De Wever *et al.* 2011). This method is similar to the one used in the module 'Mathematics in Industry', but the current model used lacks in detailed descriptions.

First year students raised concerns about the contribution level the peer assessment had in MMPS. Many felt the *'Addition to base mark'* method used needed higher weighting as it did not create enough distinction between group members when students didn't contribute evenly:

'I believe that peer assessment should be much more heavily weighted in the final marks.'
First year student

With hindsight, the second year students felt that the *'One-part weighting factor'* used in 'Mathematics in Industry', similar to the methods described by Conway *et al.* (1993), was more effective and showcased students' contribution to the group better (Conway *et al.* 1993).

The levels of confidence of all students tended towards the lower end of the scale, so time must be taken to develop how peer assessment is first presented to students. This should help them feel more confident in tackling the problem and will also increase confidence levels when transitioning year groups, as a more positive experience is had in their first involvement with peer assessment at university. For example, first years were particularly unhappy with the timing of the module, and felt that working in random groups and peer assessing early in their degrees was unnecessarily stressful:

'Don't do it in first year. Or at least, do it in the second term and let us choose our groups. It's hard to work with strangers, and getting into university life is hard enough without having to deal with peer assessment on the side.'
First year student

A reason for this could be that students do not always feel they have the skills to peer assess at this stage. However, Nulty (2011) explains that the benefits of using peer assessment early in students' university careers outweigh the detriments, as they need to gain experience through practice so that they develop these skills (Nulty, 2011).

Conclusion

Peer assessment, although a useful tool, will always be met with slight dissatisfaction from students. Past experiences can leave students feeling unconfident about undertaking peer assessment at university, and students with no experience may be apprehensive of the unknown. This, however, should not be used as a reason to remove it from modules. There is no clear divide on student opinion of peer assessment within this study, and, with research such as that by Nulty (2011) or Cheng and Warren (1997) suggesting there are more benefits to conducting peer assessment in the earlier years of degree study, it is apparent that the positive effects of peer assessment significantly outweigh the negatives perceived by students (Nulty, 2011; Cheng & Warren, 1997). This implies that peer assessment should still form a core part of the curriculum, and so consideration must be taken in how it should be implemented.

Clearly, there are benefits to both quantitative and qualitative peer assessment, therefore a combination of both will offer the most flexibility. It is relatively easy to translate a numerical assessment into a mark but this doesn't always leave students satisfied. A written component allows for more clarity as students can provide detailed reviews of their group members' contributions and staff can use these statements to quantify each individual's contribution. However, this is time consuming for staff and subjective to the marker. By combining the two methods these issues can be alleviated. The written component can be used as a method of justification for the numerical values awarded. This helps markers to understand where the values awarded come from, whilst still allowing for speedy marking where students have all worked equally well within their group.

Students wanted peer assessment to be more heavily weighted so that their individual contributions to group work was more fairly recognised. This is most easily implemented via a *'One-part weighting factor',* which gives greater precision, rather than an *'Addition to base mark'* method, the style currently used in MMPS (Conway *et al.* 1993). In the specific case of MMPS and Mathematics in Industry, a development of the peer

assessment method used in Mathematics in Industry would more closely adhere to the improvements suggested by students.

Anonymity is an area where students have a split opinion. Many second year students recognised the skills they gained from discussing with peers, although some felt unable to be critical in this situation. As a whole, first year students appreciated that the diaries were anonymous, though this can create a lot more work for staff in having to interpret the responses. In general, students are so divided that the use of anonymous peer assessment should be taken on a module-by-module basis by the module leaders. They should consider a variety of factors, including the year of study, the nature of the group work and the subject area.

Although this work has looked at peer assessment within mathematics, its emphasis has been upon approaches for students allocating marks to recognise individual contributions to group based tasks. As this is something which is, in itself, independent of the material being considered, the ideas discussed here will be of interest and relevance to those from other disciplines.

Recommendations

Peer assessment is essential for group project modules. It should:

- Encourage student ownership of their marks.
- Contain both qualitative and quantitative elements.
- Allow staff and students to easily translate marks from the raw scores.
- Make an appropriate difference to an individual student's mark depending on their contribution.
- Use anonymity with consideration of the student cohort being assessed.

Acknowledgements

Heather and Calum would like to thank Michael Grove and Rosemary Dyson for allowing us to work closely with them to build on their work in creating the original Mathematics in Industry and MMPS modules. They would also like to thank all the students who took the time to complete the questionnaires and, finally, their fellow interns for being perfect sounding boards for some of our more abstract ideas.

All authors are grateful for the project funding from the Centre for Learning and Academic Development for a project ('Enhancing Mathematics Feedback') that made this internship possible.

References

Brutus, S. & Donia, M.B.L. (2010) 'Improving the Effectiveness of Students in Groups With a Centralized Peer Evaluation System', *Academy of Management Learning & Education,* 9(4), pp652–662.

Cheng, W. & Warren, M. (1997) 'Having second thoughts: Student perceptions before and after a peer assessment exercise', *Studies in Higher Education,* 22(2), pp233–239.

Conway, R., Kember, D., Sivan, A. & Wu, M. (1993) 'Peer Assessment of an Individual's Contribution to a Group Project', *Assessment & Evaluation in Higher Education,* 18(1), pp45–56.

De Wever, B., Van Keer, H., Schellens, T. & Valcke, M. (2011) 'Assessing collaboration in a wiki: The reliability of university students' peer assessment', *The Internet and Higher Education,* 14(4), pp201–206.

Goldfinch, J. (1994) 'Further Developments in Peer Assessment of Group Projects', *Assessment & Evaluation in Higher Education,* 19(1), pp29–35.

Goldfinch, J. & Raeside, R. (1990) 'Development of a Peer Assessment Technique for Obtaining Individual Marks on a Group Project', *Assessment & Evaluation in Higher Education,* 15(3), pp210–231.

Harris, L.R. & Brown, G.T.L. (2013) 'Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: Case studies into teachers' implementation', *Teaching and Teacher Education,* 36, pp101–111.

Jacques, D. (1984) *Learning in Groups. 2nd edition.* London: Croom Helm.

McCorkle, D.E., Reardon, J., Alexander, J.F., Kling, N. D., Harris, R.C. & Iyer, R.V. (1999) 'Undergraduate Marketing Students, Group Projects, and Teamwork: The Good, the Bad, and the Ugly?', *Journal of Marketing Education,* 21(2), pp106–117.

Nulty, D.D. (2011) 'Peer and self-assessment in the first year of university', *Assessment & Evaluation in Higher Education,* 36(5), pp493–507.

Race, P. (2001) *A Briefing on Self, Peer & Group Assessment.* York: Learning and Teaching Support Network.

Webb, N.M. (1997) 'Assessing students in small collaborative groups', *Theory Into Practice,* 36(4), pp205–213.

Appendices

Appendix 1: Prompt sheet for peer assessment diaries in the first year module, Mathematical Modelling & Problem Solving.

Appendix 2: Numerical group contribution sheet used in the second year module, Mathematics in Industry.

Appendix 1

MMPS Diary Guidance

Fill in this online individual diary for project 3 which will be used (in conjunction with the diaries from the rest of your group and your Postgraduate Teaching Assistant (PGTA) feedback from the workshop sessions) to determine your contribution mark for each project. Individual comments made here will not be shared with the rest of your group, but trends across diaries and PGTA comments may be used in feedback to other group members. The deadline is the same as for your group project. You should include:

1. How your group spent their time together. Was this an effective method of working? What other approaches did you try than what ended up in report? Why were they not used? When did you meet outside of sessions?

2. How you spent your individual working time. Was this an effective method of working? What did you try? What did you take to the group and what didn't you? Why?

3. Who do you felt led on (give up to two names for each);
   a. Identifying the problem?
   b. Formulating the model (or each of the submodels)?
   c. Solving the model (or each of the submodels)?
   d. Critiquing/evaluating the model?
   e. Interpreting your solution?
   f. Writing the report?
   g. Anything else which proved necessary?

4. What went well with your project? What are you particularly proud of?

5. Which part of the project did you struggle with as a group? How did you overcome the problem? How could you improve things next time?

6. Is there some area of technical mathematics you now feel you understand better having done this project?

7. How confident did you feel tackling an unseen problem before the session started on a scale of 1–10 (1= very unconfident, 10 = very confident)?

8. How confident do you now feel in tackling an unseen problem on a scale of 1–10 (1= very unconfident, 10 = very confident)?

9. Any other comments you have.

Finally, allocate each member of your group (including yourself) a percentage which you feel reflects each person's contribution to the project. For example, if there are 5 people in your group and you think you all contributed equally then you would allocate each member 20%.

Appendix 2

Evaluating the Contribution From Each Group Member

Group Number:

Please complete the table below assigning a value for each criterion for each person using the guidance below. The figures you give will be used to calculate the Peer Assessment Factor as outlined in the assignment brief.

Student Names:					
1. Level of enthusiasm/participation					
2. Suggesting ideas					
3. Understanding what was required					
4. Helping the group to function well as a team					
5. Organising the group and ensuring things got done					
6. Performing tasks efficiently					

For each criterion marks are awarded as follows:

3 for *'better than most of the group in this respect'*
2 for *'about average for this group in this respect'*
1 for *'not as good as most of the group in this respect'*
0 for *'no help at all in this respect'*

Signatures

Name:		Signature:	
Name:		Signature:	
Name:		Signature:	
Name:		Signature:	
Name:		Signature:	