Group work and Fair Assessment: A Case Study

Dale Parsons
Diana Kassabova

Otago Polytechnic
Dunedin, New Zealand
dale@tekotago.ac.nz

[2]Auckland University of Technology

Proceedings of the 15th Annual NACCQ, Hamilton New Zealand July, 2002 www.naccq.ac.nz

ABSTRACT

There are sound educational and vocation reasons for using group work in higher education. Employers value teamwork skills and as part of professional practice a graduate needs to be able to critically evaluate their own work (self-assessment) and that of their peers. It is also important that students learn about their effectiveness in a group setting (Lejk, Wyvill and Farrow, 1996; Boud, Cohen and Sampson, 1999; Rafiq and Fullerton, 1996).

Educationally group work is seen as providing a vehicle through which students can be involved in deep learning, developing their skills experientially and contributing to the skills they will need for life long learning (Bourner, Hughes and Bourner, 2001). Group work emphasises co-operation over competition and can promote respect for other group members' experiences and values (Boud *et al.* 1999).

However, the assessing of group work can be problematic for both students and lecturers (Lejk, *et al.* 1997). When all students in a group are awarded the same mark, many students have experienced the problem of 'passengers' or 'freeloaders' (Bourner *et al.* 2001). Often students feel this mark is not a fair reflection of each group member's contribution to the piece of work (Conway, Kember, Sivan and Wu, 1993). The problem areas for lecturers are in assessing whether all students have contributed equally and on how valid and reliable a group mark is.

1. A CASE STUDY: A CATEGORY BASED APPROACH

In the BIT programme at Otago Polytechnic there have been cases where members of a group have contributed much more to the group assignment than others but because of the assessment method used have ended up being awarded the same mark. In one instance, the least able student in the class teamed up with one of the top students. They received one of the top marks for their group work although from the assignment report it was easy to conclude that the top student had done all the work. The marks were allocated equally amongst the group members and there were no protests about poorly performing group members or dysfunctional groups until after the assessment event.

This case raised a concern and triggered an interest in exploring alternative ways for assessing group work. It was decided that a pilot study would be carried out that would introduce a new method for group assessment to the BIT programme.

2. THE GROUP ASSESSMENT METHOD

An excellent survey of group assessment methods can be found in the paper by Lejk *et al.* 1996. The method we chose involved the lecturer assessing the final product produced by each group and the students peer- and self- assessing the process of working in the group. The students would end up generating a plus or minus contribution mark for themselves and each member of their group. An individual student would receive the product mark modified by the average of the peer marks allocated to them. The ultimate aim was to acknowledge and reward or penalise individuals' contributions to the group work without introducing competition within the group.

3. DESCRIPTION OF THE GROUP ASSIGNMENT

The course was a year two module in the Bachelor of Information Technology degree. There were four assessments in the course: two theory tests and two group assignments. The weightings for the course are shown in Table 1.

Assessment	Weight
Assignment 1	0.25
Assignment 2	0.15
Test 1	0.3
Test 2	0.3

Table 1: Assessment weightings

The pilot study was carried out on the second assignment. It was decided that as its weight was relatively low (0.15 of the total mark for the paper) any possible problems that might arise with the new assessment method would have a minimum effect on the final marks. For this assignment the student groups were given a list of topics related to different aspects of the Internet to choose from.

The aim of the assignment was to get the students to research a topic and then relate their findings to a framework that they had previously examined in class. Each group had to submit a research report on the topic selected by the group and also present their topic to the class.

Students in each group could receive different marks depending on their contribution. There were 46 students in the course divided into 16 groups.

Relative contribution to the project					
Suggested criteria for peer assessment	Didn't contribute in this way	Willing but not very successful	Average	Above average	Outstanding
Researching the topic	-2%	-1%	0%	1%	2%
Report writing	-2%	-1%	0%	1%	2%
Ideas and suggestions	-2%	-1%	0%	1%	2%
Presentation preparation	-2%	-1%	0%	1%	2%
Referencing and handouts	-2%	-1%	0%	1%	2%
Preparing demonstrations	-2%	-1%	0%	1%	2%

Table 2: Descriptors for the category based student contribution scoring system

Assessor's name:				
Rita	Rita	Perry	Jack	Total
Researching the topic	0	0	0	0
Report writing	0.5	0.5	-1	0
Ideas and suggestions	0	0	0	0
Presentation preparation	0	0	0	0
Referencing and handouts	1	1	-2	0
Preparing demonstrations	0	0	0	0
Total:	1.5	1.5	-3	0

Assessor's name				
Perry	Rita	Perry	Jack	Total
Researching the topic	0.5	0.5	-1	0
Report writing	0.5	0.5	-1	0
Ideas and suggestions	0	0	0	0
Presentation preparation	0	0	0	0
Referencing and handouts	0.5	0.5	-1	0
Preparing demonstrations	0	0	0	0
Total:	1.5	1.5	-3	0

Assessor's name:				
Jack	Rita	Perry	Jack	Total
Researching the topic	0	0	0	0
Report writing	1	1	-2	0
Ideas and suggestions	0	0	0	0
Presentation preparation	0	0	0	0
Referencing and handouts	0	0	0	0
Preparing demonstrations	0	0	0	0
Total:	1	1	-2	0

Table 3: Self and Peer Assessment example

Stream:	**C**		
Topic:	**WWW**		
Students names:			
Perry, Jack, Rita	**Max marks**	**Marks obtained**	**Comments**
Report			
Overview of the topic	**5**	**2**	You have given a **historical** overview of the Web but not an overview of your topic
Main points/concepts explained and supported by examples	**25**	**20**	There is no indication of where the HTTP protocol fits in the OSI or TCP/IP models; There is also some repetitiveness when explaining concepts;
Conclusion and future directions	**10**	**3**	your conclusion looks more like a summary
Layout and structure	**10**	**8**	more care with grammar is needed
Reference style	**10**	**10**	well done!
Handouts for the class	**10**	**9**	
Presentation			
Presentation style	**15**	**12**	some concepts (e.g. sockets) were not explained well
Visuals (relevant to the topic**)**	**10**	**7**	more visuals would've made it more understandable
Timing	**5**	**5**	
Total	**100**	**76**	

Students names:	**Ave Peer mark**	**Final mark**
Perry	1	**77**
Jack	-3	**73**
Rita	1	**77**

Table 4: Lecturer feedback example

Incorrect or all members did not submit marks	All zeros	Differentiated-group in agreement	Differentiated – disagreement within the group
3	8	1	4

Table 5: Inter-rater Consistency

4. IMPLEMENTING THE METHOD

The lecturer discussed at length with the students the purpose of the assignment and why there was an emphasis on group work. Important points for working effectively in a team were discussed and the students reportedly found this helpful.

Students were given the opportunity to select their own group. Although this approach does not necessarily simulate real life situations, it was applied with the aim to facilitate a positive attitude towards group work and to give the students a feeling of control.

Students received:

- a copy of the assignment
- the lecturer's marking schedule
- a list of suggested peer assessment criteria and marking scale
- a worked example showing peer marking and the final score per student.

In order to make very clear what was required all the above documents were given out and explained in detail at the same time. Each group were asked to discuss and possibly change the suggested criteria and submit their group's set of criteria before they embarked on working on the assignment.

After they had submitted the research report and given the presentation, the students submitted their self and peer assessment forms. Students scored each other on a scale of -2 to 2 for each of six categories. A score of zero meant that the student had made an average contribution for this category. Table 2 shows the scale descriptors.

Each student submitted their table separately and they did not see what marks their peers allocated to them. Table 3 presents one group's self and peer assessment. This group is basically in agreement as to which student has contributed the most and least to the assignment. Their order of students is the same; one student's marks have a lesser range than the other two.

Table 4 is an example of the lecturer feedback and final mark allocated for each student. The assessment of the product and the peer assessment of the process can be clearly seen as well as the final mark for each group member.

Figure 1 shows the peer marks received by all the students. The range was from -4 to 3 with most groups opting to give each other zero which meant that all members of the group received the same final mark. The raw peer marks ranged from -6 to 6, so averaging the group marks has had a moderating effect on the mark.

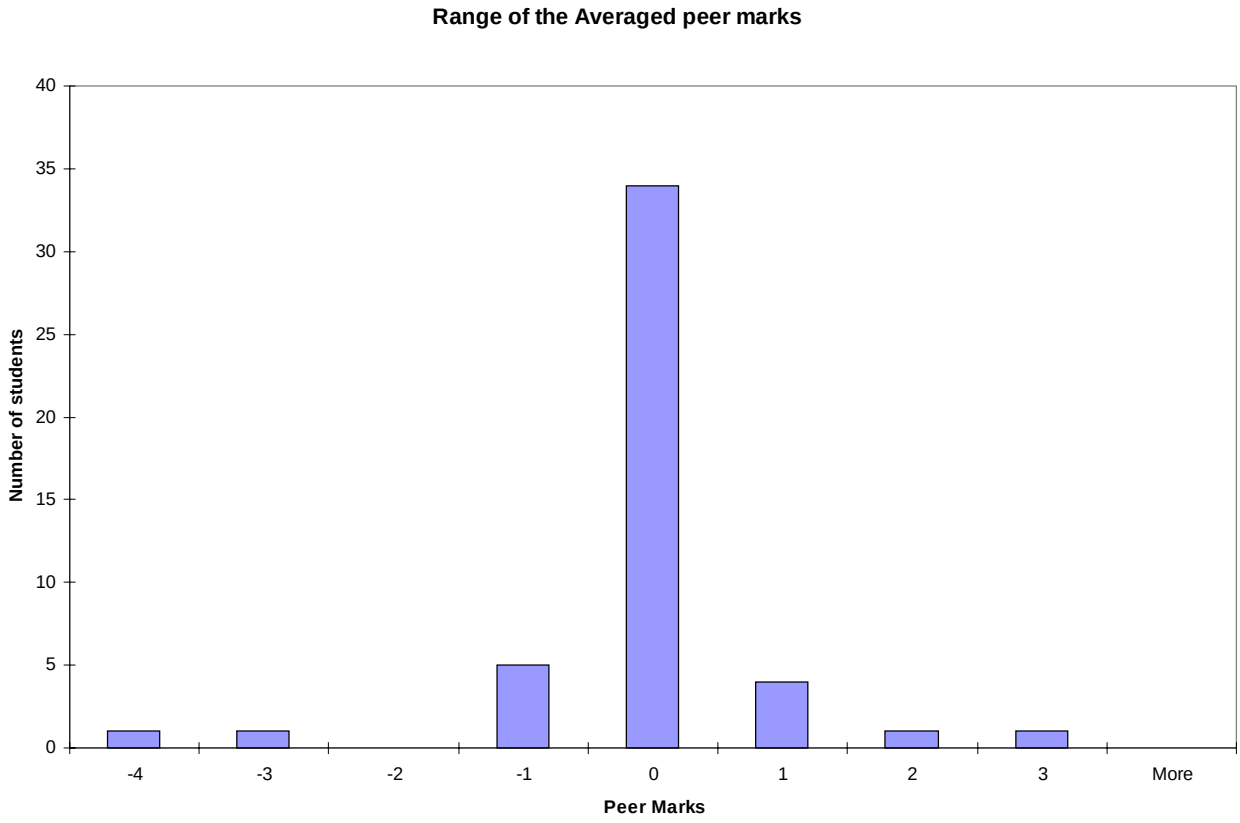
5. CRITICAL ANALYSIS

In order to critically analyse the assessment practice used for this study, certain criteria should be taken into account: validity, reliability of the practice, fairness, integration of the assessment into the teaching/learning programme, and manageability of the marking.

Generally, validity of the assessment could be affected if the scoring fails to capture important qualities of task performance. In this study the marking schedule was very clear and well aligned to the tasks, therefore this was not considered an issue. The suggested criteria for peer marking focussed on the process of working in the group and on individual contributions to the work. Some groups used it as a way to split up the tasks; this would indicate that the criteria are closely aligned with task performance. The students were also able to create their own criteria. Three groups out of the 16 made changes to the suggested criteria.

Another aspect of validity for this type of assessment is inter-rater consistency, i.e. whether group members scored each other in a similar manner. If they were marking fairly they should have been in agreement as to who was the greatest contributor. Table 5 shows inter-rater consistency for the case study. The first category shows the number of groups where group members did not follow marking instructions or did not hand their marks in. The 'All zeros' category shows the number of groups where everyone in the group scored the other members as average contributors and therefore all members received the same group mark. The third category was where the group awarded members' different marks based on their differing contributions and the group was in agreement. The last category shows the number of groups that did not score consistently between members.

Two threats to validity arise out of this table: one is what the lecturer did with the marks and the other is what the students in the disagreement groups did. The lecturer gave the 3 incorrect groups equal group marks. Perhaps it would have been better if she had gone back to the groups and got them to correct or submit missing marks. The four inconsistent groups received averaged peer marks.



In this situation the lecturer could have replied back to the groups and told them that their marks are inconsistent, and asked them to reconsider.

This raises the question as to whether the peer marks should be anonymous. Perhaps in a disagreement situation they should not be and the students should have to negotiate with each other until they reach a form of agreement.

The way the assignment was organised is considered to have led to worthwhile benefits to the students as they gained knowledge, collaboration skills and credit towards their final result. From observing students' work on the assignment and from the quality of the final product the lecturer came to the conclusion that the group work was help rather than a hindrance to the motivation of the students. However, there was one group that identified themselves as not being able to work together. This must have resulted in a negative impact on their relationships and their motivation.

One of the major reasons for introducing this assessment method was to achieve a greater degree of fairness in the marking of group work. For different individual contributions this method redistributed some of the marks. The marks allocated by the lecturer were based on criteria therefore the students were judged on their performances in relation to the criteria rather than in comparison to others; this according to Gipps (1994) is fairer than judging students on a norm referenced model.

The assignment was well integrated into the teaching and learning programme. Being the last piece of work for this paper it acted as a bridging exercise towards the third year networking paper.

6. CONCLUSION

Is the plus or minus contribution method a good way of rewarding individual contributions to a project? The advantage for the lecturer is one product to mark per group. For the students they have the possibility of being rewarded for a greater contribution to the group.

There were several problems with this assessment method. The major problem was manageability. The lecturer received 43 tables. Nearly 20% of the groups had members who either submitted incorrect tables or did not submit a table. It was found that the amount

of time to collect, collate and integrate the peer marking with the lecturer's product mark was unacceptably high. This task needs to be simplified or automated.

A weakness that was obvious in some groups was the lack of inter-rater consistency, i.e. some groups were in disagreement as to whose contribution was the highest and the lowest. Further consideration should be given to ways of dealing with such cases.

In this case study the range of marks that were awarded by the student contribution component was very low (-4 to +3) with half the groups agreeing to an equal sharing of the group mark. This could have meant that all the members of each group were contributing equally or that the students were marking very conservatively because they were not used to assessing each other.

7. DIRECTIONS FOR FURTHER RESEARCH

We have decided to move away from a category based peer assessment approach to a more holistic approach due to the manageability problems we experienced. Other doubts about the use of the category based approach were raised in a paper by Lejk and Wyvill (2001). They also point out that most groups divide the workload so it is not appropriate to score each student on every category. These scoring systems frequently imply that all categories are equally important when in fact this may not be the case.

The holistic approach that we are going to trial involves the students giving each other one mark. The students will still develop marking criteria, which they will use in formulating this holistic mark. The Lecturer will then work out each student's mark based on his or her individual contribution mark and the group product mark.

REFERENCES

Boud, D., Cohen, R. & Sampson, J. (1999). Peer Learning and Assessment. Assessment & Evaluation in Higher Education, 24(4), pp. 413-426.

Bourner, J., Hughes, M. & Bourner, T. (2001). First-year Undergraduate Experiences of Group Project work. Assessment & Evaluation in Higher Education, 26(1), pp. 19-39.

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), pp. 45-54.

Gipps, C. (1994). Developments in Educational Assessment: what makes a good test? Assessment in Education, 1(3), pp. 283-291.

Lejk, M., & Wyvill, M. (2001). Peer Assessment of Contributions to a Group Project: a comparison of holistic and category-based approaches. Assessment & Evaluation in Higher Education, 26(1), pp. 61-72.

Lejk, M., Wyvill, M., Farrow, S. (1996). A Survey of Methods of Deriving Individual Grades from Group Assessments. Assessment & Evaluation in Higher Education, 21(3), pp. 267-280.

Lejk, M., Wyvill, M., Farrow, S. (1997). Group Learning and Group Assessment on Undergraduate Computing Courses in Higher Education in the UK: results of a survey. Assessment & Evaluation in Higher Education, 22(1), pp. 81-91.

Rafiq, Y. & Fullerton, H. (1996). Peer Assessment of Group Projects in Civil Engineering. Assessment & Evaluation in Higher Education, 21(1), pp. 69-81.