

Making a Difference: Using Peers to Assess Individual Students' Contributions to a Group Project

Abstract Overcoming the potential dilemma of awarding the same grade to a group of students for group work assignments, regardless of the contribution made by each group member, is a problem facing teachers who ask their students to work collaboratively together on assessed group tasks. In this paper, we report on the procedures to factor in the contributions of individual group members engaged in an integrated group project using peer assessment procedures. Our findings demonstrate that the method we used resulted in a substantially wider spread of marks being given to individual students. Almost every student was awarded a numerical score which was higher or lower than a simple group project mark would have been. When these numerical scores were converted into the final letter grades, approximately one third of the students received a grade for the group project that was different from the grade that they would have received if the same grade had been awarded to all group members. Based on these preliminary findings we conclude that peer assessment can be usefully and meaningfully employed to factor individual contributions into the grades awarded to students engaged in collaborative group work.

Background

The use of group work in higher education has become commonplace for sound pedagogical reasons. These reasons include facilitating learning and working as part of a team which is essential in modern industry and business (Jacques, 1984); developing interpersonal relations and individual responsibilities (Oldfield and MacAlpine, 1995:

126); and developing 'personal transferable skills' of communication, presentation, problem-solving, leadership, delegation and organization (Butcher *et al.*, 1995:165). In addition, Butcher *et al.* (1995: 165) argue that group work can provide a more interesting, effective and hence preferred learning context compared to traditional lectures and help to facilitate learning and teaching in very large classes. Other benefits of group work include the co-construction of knowledge which concurs with Vygotsky's (1978 and 1981) claim that interactions produce new understandings for the participants. Group work also promotes learning through encouraging discussions and debate which in turn encourage the justification of ideas, resolution of disagreements and understanding of new perspectives (Webb, 1995: 244).

While those who have studied group work operating in practice list a number of advantages to this form of collaborative learning, group work is not without its drawbacks. One disadvantage of group work is 'social loafing' or 'free riding' as described by Webb (1995: 245-246) which takes place when a member of a group contributes little or nothing to the work of the group. The effect of this can be to demotivate other group members who resent doing all the work and so the efforts of the entire group spiral downwards. Webb (1995: 246-247) also notes that the division of labour, which may take place within the group, may not be beneficial in terms of individual learning because students only become involved in part(s) of the group work and so might be denied a sense of completeness.

Another possible drawback is that group work may present problems if it is to be formally assessed and graded. For as Butcher *et al.* (1995: 165) point out, group work raises the question of whether or not all the students in the same group should receive the

same mark or whether they should be awarded individual marks reflecting the contribution of each student. This issue is central to the research presented in this paper and is examined in more detail in the following sections.

Assessing Group Work

Gibbs, Habeshaw and Habeshaw (1986) present evidence that group work potentially creates assessment problems both within and across groups. Within a group, students might be unhappy with the same mark being given for all group members. Between groups, there is often a narrower and hence less discriminating range of marks, resulting from group marks rather than individual marks being given. Gibbs *et al.* (1986) also point out that group work often results in very good work being produced for which grades are both higher than usual and in a narrower range.

Group work is very often criticised as being inequitable when the same assessment is given to all group members irrespective of the contribution or efforts individual members of the group have put in (see for example, Goldfinch and Raeside, 1990; Conway *et al.*, 1993). This common criticism of group work emanates from both teachers and students. The problems associated with awarding a group mark or grade may take a variety of forms. The unfairness might be to the benefit of group members who 'free load' (Webb, 1995: 246) and subsequently are awarded a grade for which they did little or no work. Alternatively, it could be to the benefit of those who have made an effort but whose quality of work input still falls short of the overall quality of the final group product and the grade awarded. Finally, it can be unfair to the students whose contributions in terms

of both effort and input have amounted to more than the mark or grade awarded to the total group work.

What then can be done to overcome these potential problems and to ensure that the mark or grade awarded to a group member more fairly and accurately reflects the respective contributions he or she has made to the final product? One way is to actively and purposefully involve students in assessing the contributions of their peers in the process of group work. This paper describes the use of peer assessment of group members' individual contributions as a method to facilitate the differentiation of assessments in the context of group work. The impact of this method on both the numerical scores and the final grades awarded to students is analysed and the implications of the findings are discussed.

Peer Assessment

Falchikov (1986: 146) notes that the 'prevailing model of assessment in higher education is authoritarian' in which the students are effectively excluded from every stage and the possession and exercise of power is unequal. This assessment method goes against the development of students' responsibility and autonomy (Falchikov, 1986: 146-147). Falchikov and Magin (1997: 386) also observe that the system is unreliable as 'correlations between scores of markers are low', resulting in 'a measure of injustice'. Peer assessment is thus perceived as one means of redressing these problems which are often associated with traditional methods of assessment. The practice of peer assessment brings a number of benefits which earlier studies have documented. For example, peer assessment develops responsibility and enterprise (Goldfinch and Raeside, 1990: 210)

and increases students' maturity and confidence. Peer assessment acts as a socialising force and enhances relationships and relevant skills (Earl, 1986: 68). Brindley and Scoffield (1998: 88) have found that deep learning occurs with students reporting that they perform better and have been exposed to new ideas as a direct result of participating in peer assessment. Peer assessment also promotes autonomous learning, reflective learning and less dependence on the teacher as the supposed expert (Brindley and Scoffield, 1998: 88). Falchikov (1986: 161) reports that students think more, learn more, and become more critical as a result of participating in peer assessment. Another advantage of peer assessment that is of particular interest in this study is that it can be used as a means of individualising students' grades in a collaborative learning context. In this way, the kind of free riding problem or unfairness of uniform grading within a group described earlier can be discouraged by awarding students marks or grades based on their individual contributions to the group work. This can often best be achieved through the peers assessing one another's contributions and it is this aspect of peer assessment which is examined in this paper.

Peer Assessment of Group Work

In a group work situation, having peers to assess one another's contributions has been advocated to help to overcome the kinds of problems characteristic of group work discussed earlier (for example, Conway *et al* , 1993; Goldfinch and Raeside, 1990; Kane and Lawler, 1978; Falchikov, 1986; and Falchikov and Magin, 1997). Students are in a unique position to assess their peers in terms of individual contributions or effort made which is, in most cases, entirely inaccessible to their teacher(s). However, it is exactly

due to this that this practice has been criticised in some studies as lacking reliability (see for example, Boud, 1986; Swanson *et al.*, 1991). A study by Falchikov and Magin (1997: 386) suggests that low rater reliability can be overcome with the use of multiple ratings and this claim is in line with other studies (see for example, Fagot, 1991; Houston *et al.*, 1991; Magin, 1993). In another study, Falchikov (1986) has shown that devolving the assessment of group processes to peers can be carried out with a reasonable degree of reliability although peer-teacher correlational analysis is obviously not possible in situations where only students give assessments.

Different ways of distinguishing between the performance of individuals in a group have been attempted by researchers and educators. These studies include each student submitting a portion of the group work and the teacher sitting in the group meetings to decide on individual contributions (Earl, 1986). In theory at least, in a group context, peer assessment allows both process and product of learning to be assessed more fairly and accurately because it is only the students themselves who have knowledge about and access to all of their group's activities and hence the contributions of individual group members to the entire group work. Earlier studies have been promising and the individual marks are considered reflective of students' merits and contributions (see for example, Goldfinch and Raeside, 1990: 222). Peers can be asked to assess the contributions of their peers basically in one of two ways. The first option is to ask the group to openly discuss and negotiate among themselves the contribution of each group member. This method may be problematic, however, when the aim is to differentiate between the grades awarded to the group members if the students find it difficult to assess one another in such a public forum. The second option is to ask individual group

members to assess one another anonymously by using a peer assessment form. This method lacks the openness of the first option, but it overcomes the possibility of bad feelings or the problem of students being mindful of offending their peers when assessing their contributions. Conway *et al.* (1993), for example, devise a form for students to complete confidentially and this practice was followed in this study.

The Present Study

In this study, we have adopted the procedure advocated by Conway *et al.* (1993) as an improvement on the model they used in their study, but which, to our knowledge, has never been examined as to whether or not it makes a significant difference to the marks awarded; and that is, between a group mark for all and individual marks based on individual contributions. The procedure involves factoring in of individual group members' effort or contributions to the group project with a view to differentiating assessments. In determining the success of this procedure, we have been mindful of Heron's (1981) point that an assessment procedure should be 'valid, reliable, practicable and fair, and useful to the students'.

Methodology

Three classes, comprising 16, 18, and 19 Hong Kong Chinese students (N = 53), all of whom were first year BSc (Hons) Electrical Engineering undergraduates studying the course English for Academic Purposes (EAP), took part in the peer assessment exercise. Halfway through the course, the students formed themselves into groups of 4 or 5 to carry out an integrated group project. The group project was made up of three components: a

seminar presentation, an oral presentation, and a written report. Each component was carried out as a collaborative group activity. The project was an assessed assignment within EAP and contributed 40% towards the final course grade given to each student at the end of the one semester course. The assessment procedures included both teacher assessment and peer assessment with each carrying 50% of the assessment weighting. The assessment criteria for each of the group project components (see Appendix) had been devised, trialled and found to be satisfactory by teachers responsible for assessment of the course within the department.

To the students, group work was not a new experience. In a number of their courses at the university, they are required to work in groups. For example, in the EAP class, before embarking on the group project, the students had a lot of opportunities to perform various small tasks in groups, and they had been given explicit instruction in group discussion skills. Peer assessment, however, was a new experience for the majority of the students. In order to ensure the reliability and integrity of the peer assessments, in the lead up to the actual assessment of the three project components, time was set aside to orient the students to the notion of peer assessment, and to give them opportunities to discuss the assessment procedures, especially the assessment criteria for the project components in terms of their importance, meaning and use, as well as to practise assessing one another's. Care was taken to make this orientation process uniform across the three classes. Students were also given opportunities to practise assessing their peers earlier in the course and these assessments were evaluated by both teachers and students and proved to be very useful.

At the end of the group project, each group was awarded a Final Group Project Mark which was determined by a combination of marks given by both the teacher (50%) and the rest of the class (50%). Each of the three components of the group project was given a separate numerical score, and so this Final Group Project Mark was a compilation of six sets of scores.

Calculating Individual Contributions

In order to more realistically and objectively reflect individual contributions to the group project, each student in the group was awarded an individual mark, rather than all the group members being awarded the same mark. This study chose to basically adopt Conway *et al.*'s 'simpler scheme with a one-part multiplicative weighting factor' (1993:45). In this scheme, the final mark awarded to the individual student is the product of the Final Group Project Mark and the Individual's Weighting Factor (IWF) as illustrated in the formula below (Conway *et al.*, 1993: 50):

$$\text{Final Individual Student Mark} = \text{IWF} \times \text{Final Group Project Mark}$$

The IWF was specifically devised to overcome the possible unfairness of giving all members of a group the same grade regardless of differing contributions from individual students. The Individual Weighting Factor was derived from the following formula (Conway *et al.*, 1993: 50):

$$\text{Individual Weighting Factor} = \frac{\text{Individual Effort Rating}}{\text{Average Effort Rating for Group}}$$

The Individual Effort Rating was obtained by summing the marks for seminar presentation, oral presentation, and report awarded to an individual student by the rest of

the group. The Average Effort Rating for each group was obtained by summing all Individual Effort Ratings for a group and then dividing the total by the number of group members. The actual mark awarded to an individual student was hence obtained by weighting the Final Group Project Mark up or down by factoring in the contribution or effort rating (IWF). By using this method, individual students were rewarded according to the percentage of the group's success to which they had contributed.

Assessing individual effort within the group was achieved in the following way. Following the completion of the group project, each member of the group was asked to rate the effort or contributions of the other group members to six specified tasks relating to the group project on a six-point scale (see Table 1).

(INSERT Table 1)

In all the preceding assessments of the various components of the group project, the teacher and the students had shared the awarding of marks equally, but the students alone assessed the contributions of their peers which in turn determined the IWF. This element of the peer assessment is therefore unique in that only the students themselves could award marks for this important part of the group process, i.e. the process.

Results

Table 2 presents the numerical scores, or marks, for each of the students in the study without and with the IWF. In other words, the numerical scores are listed as group marks and as individual marks. These have then been converted into letter grades enabling

comparisons to be drawn between the grades that would have been awarded if members of each group had been given the same grade and the individual student grades that were in fact awarded.

(INSERT Table 2)

In Table 2, it can be seen that the IWF was greater or less than 1 for 50 of the 53 students' numerical scores, and so it impacted 94% of the students. The IWF ranged between 1.1 and 0.857 and had a negative effect on students' numerical scores in 29 (54.7%) instances and a positive effect in 21 (39.6%) cases. Nil effect was observed in three of the 53 cases. Before the IWF was factored in, the group marks ranged between 56.63% and 68.09% with a spread of 11.54 percentage points between the highest and the lowest scores. The application of IWF was found to have a marked effect on the range of the numerical scores. With the IWF, the scores ranged between 48.1% and 74.21% with a spread of 25.4 points between the top and the bottom scores. The IWF can be seen to have more than doubled the spread of grades from 11.54 to 25.4 percentage points; and in so doing, it has pushed the lowest score down by 7.82 points and raised the highest score by 6.11.

Interestingly, the result of a paired *t*-test of the differences between group marks and individual student marks (see Table 3) shows that the difference between the means of the two sets of marks was very small and not statistically significant at the 0.05 level.

(INSERT Table 3)

Table 3 shows that the individual student marks had a larger standard deviation than the group marks which supports the results reported above. While there is no doubting the impact of the IWF on numerical scores, what interests us most was whether this impact was sufficient to effect any differences in the final letter grades awarded to individual students for the total group project. The letter grades used indicate a band of numerical scores determined by our university and are detailed in Table 4.

(INSERT Table 4)

Altogether there are 9 grades. Grades in the range A+ to D are passing grades and grades E and F are both failing grades. With the exceptions of the top (A+) and bottom (F) grades, each letter grade represents 5 percentage points and so a change in a student's numerical score does not necessarily translate into a change in her or his letter grade. Table 5 shows the number of different grades given to students both before and after the IWF was factored into the numerical scores.

(INSERT Table 5)

As one might expect, the impact of the IWF on the final grade given to the students was less than it had been on their numerical scores because of the fact that each letter grade represents at least a band of five percentage points. While almost every numerical score changed as a result of the IWF, the same effect was not observed when these scores were converted into the final letter grades. Nevertheless, 17 (32%) of the 53 students received a different grade as a result of factoring in the IWF (shown in bold italics in Table 2). The overall impact of factoring in the IWF was that the final grades, instead of bunching across three grades (C, C+ and B), now fell across six grades (i.e. B+ to E). The movement in grades awarded was in both directions with seven receiving a higher grade and ten a lower grade. In most cases (16 out of 17), the movement was a difference of one grade, but for one student (student no. 32) the movement was two grades. For this student the final grade moved from C (satisfactory) down two grades to E (fail), a significant difference for the individual concerned.

Conclusions

A number of peer assessment studies cite Conway *et al*'s (1993:45) 'simpler scheme with a one-part multiplicative weighting factor' when awarding assessments to individual members in collaborative group work, and yet not until this study has the actual impact of Conway *et al*'s method on measuring and registering individual efforts been closely examined. The major conclusion of our study is that an IWF determined by peers definitely has a clear impact on the numerical scores attained by individual students in that the scores can much better and more fairly describe and differentiate the contributions of individual students in their group projects.

For practical reasons, numerical scores were to be converted to letter grades to conform to the university-wide assessment conventions. Our concern, as both teachers and researchers, was whether this impact was sufficient to affect any differences in the final letter grades awarded to individual students. Though *t*-test results were not significant, a change in grades for one third of the students has made this a valid and meaningful procedure.

There are plans to integrate the procedure of assessing peers' contributions into group work tasks for other classes, and to recommend this procedure to colleagues as well. A much more thorough and comprehensive orientation and training programme, which includes the elements of awareness raising, practice, evaluation and reflection, would be organised for both students and teachers. All students should understand and acknowledge the purpose, importance and usefulness of the procedure so that they know the expectations of them as members participating in group work and how they should contribute towards collaborative team work. Another reason that orientation and training are essential is for the assessment procedure to be fairly, objectively and usefully implemented. Concerning the criteria on which to assess individual student's efforts in the process of the group work, it would be more useful if students are involved in designing and writing them. Input from both the teacher and students will help to construct a list of assessment criteria specific to both the particular group work task and the particular student composition. Evaluation and self- and group-reflection at the completion of the peer assessment procedure is beneficial, not only from the point of view of course evaluation and future pedagogic design, but most important of all, for enriching the learning experience of the students.

This study, in which all of the subjects were Hong Kong Chinese, did not examine the role of cross-cultural factors in the peer assessment of group work. Future studies might seek to identify and describe such factors by drawing comparisons with students from different cultural backgrounds.

The logistics of conducting peer assessment are more convoluted using this system. With refinement of the system and use of suitable software, we believe that the method could be made simpler and less time consuming to implement. There is no doubt that involving students in making choices and evaluations during their learning process is beneficial. It is also important that students should be involved in assessing one another's, not just performance, but efforts or contributions put into the learning process. An equally important objective is to alert the students to the expectations both of the teacher and their group members in terms of their contributions to team work even before they embark on the group assignment.

This method of determining individual marks for students submitting group work facilitates the benefits of group work while providing opportunities for peer assessment. Peer assessment of group member's contributions proved to be a realistic and realisable objective which served to also enhance relationships within and between students groups and promote skills relevant to their future professional needs.

Acknowledgements

The work described in this paper was substantially supported by a grant from the Research Grants Council of the Hong Kong Special Administration Region (Project No.

G-S252). We are grateful to Ronald Chan for his assistance in preparing the statistical data.

References

- BOUD, David J. (1986) *Implementing Student Self-assessment*. Sydney: Higher Education Research and Development Society of Australia.
- BRINDLEY, Clare and SCOFFIELD, Susan (1998) Peer assessment in undergraduate programmes. *Teaching in Higher Education*, 3/1, pp. 79-89.
- BUTCHER, A. C., STEFANI, Lorraine A. J. and TARIQ, Vicki N. (1995) Analysis of peer-, self- and staff-assessment in group project work. *Assessment in Education*, 2/2, pp. 165-185.
- CONWAY, Robert, KEMBER, David, SIVAN, Atara and WU, May (1993) Peer Assessment of an individual's contribution to a group project. *Assessment and Evaluation in Higher Education*, 18/1, pp. 45-56.
- EARL, Shirley E. (1986) Staff and Peer Assessment – Measuring an individual's contribution to group performance. *Assessment and Evaluation in Higher Education*, 11/1, pp.60-69.
- FAGOT, Robert F. (1991) Reliability of ratings for multiple judges: intraclass correlation and metric scales. *Applied Psychological Measurement*, 15/1, pp. 1-11.
- FALCHIKOV, Nancy (1986) Product comparisons and process benefits of collaborative peer group and self assessments. *Assessment and Evaluation in Higher Education*, 11/2, pp.146-166.
- FALCHIKOV, Nancy and MAGIN, Douglas (1997) Detecting Gender Bias in Peer

- Marking of Students' Group Process Work. *Assessment and Evaluation in Higher Education*, 22/4, pp. 385-396.
- HOUSTON, Walter, RAYMOND, Mark R. and SVEC, Joseph C. (1991) Adjustment for rater effects in performance assessment. *Applied Psychological Measurement*, 15/4, pp. 409-421.
- GIBBS, Graham, HABESHAW, Sue and HABESHAW, Trevor (1986) *Interesting Ways to Assess Your Students*. Bristol: Technical and Educational Services Ltd.
- GOLDFINCH, Judy and RAESIDE, Robert (1990) Development of a Peer Assessment Technique for Obtaining Individual Marks on a Group Project. *Assessment and Evaluation in Higher Education*, 15/3, pp. 210-231.
- HERON, John (1981) Assessment revisited. In D. Boud (ed.) *Developing Student Autonomy in Learning*, London: Kogan Page.
- KANE, Jeffrey S. and LAWLER, Edward E. (1978) Methods of Peer Assessment. *Psychological Bulletin*, 85/3: 555-586.
- JAQUES, David (1984) *Learning in Groups*, London: Croom Helm.
- MAGIN, Douglas (1993) Should student peer ratings be used as part of summative assessment? *Research and Development in Higher Education*, 16, pp. 537-542.
- OLDFIELD, Keith A. and MacALPINE, J. Mark K. (1995) Peer and self-assessment at the tertiary level – an experiential report. *Assessment and Evaluation in Higher Education*, 20/1, pp. 125-132.
- SWANSON, David B., CASE, Susan M. and Van Der VLEUTEN, Cees P. M. (1991) Strategies for student assessment. In David J. BOUD and Grahame FELLETTI (eds.) *The Challenge of Problem Based Learning*. London: Kogan Page, pp. 260-

273.

VYGOTSKY, Lev S. (1978) *Mind in society: the development of higher psychological processes*. Cambridge MA: Harvard University Press.

VYGOTSKY, Lev. S. (1981) the genesis of higher mental functioning. In J.V. WERTSCH (ed.) *The concept of activity in Soviet psychology*. Armonk: New York, pp. 144-188.

WEBB, Noreen M. (1995) Group Collaboration in Assessment: Multiple Objectives, Processes, and Outcomes, *Educational Evaluation and Policy Analysis*, 17/2, pp. 239-261.

APPENDIX

Assessment Criteria for the Project Components

The following rubric and scale were used for each of the three group project components:

Assessing your peers is not an easy task. You need to try to be fair and objective. Your assessment scores will only be seen by your teacher.

Use the following scale when assessing your fellow students

| | | | | |
|------|------------------|---------|------------------|-----------|
| 1 | 2 | 3 | 4 | 5 |
| poor | below average | average | above average | excellent |

Seminar Assessment Criteria

A. Preparation, Overall Presentation and Content of Seminar Paper

- evidence of rehearsal
- relevance and interest of material for the audience
- quality and appropriacy of visual aids
- well-structured with clear conclusion(s)

B. Delivery of Seminar Paper

- rapport with and sensitivity to audience
- body language, timing and pacing
- use of visual aids
- clarity and coordination of group delivery
- encouragement of discussion

- C. Participating in Seminar Discussion**
 - relevant questions, comments, ideas
 - appropriate academic language
- D. Language**
 - accurate and appropriate use of vocabulary, structures and register
 - concise and clear expression of ideas

Written Report Assessment Criteria

- A. Preparation**
 - clear objective/purpose statement
 - appropriate methods/procedures for collecting information/data
 - selection of information relevant to topic and purpose
 - organisation of parts of report relevant to objective and reader interest and needs
- B. Organisation**
 - good connection of ideas
 - appropriate use of a variety of cohesive devices
- C. Language**
 - accurate and appropriate use of vocabulary, structures and register
 - concise and clear expression of ideas
- D. Writing style**
 - objective, concrete and organised facts
 - appropriate use of reference conventions
- E. Layout and presentation**
 - attractive and appropriate use of paragraphing, headings, numbering, spacing and illustrations

Oral Presentation Assessment Criteria

- A. Quality of preparation**
 - relevance and interest value of topic for audience
 - structuring of presentation
 - quality of content
 - quality and appropriacy of visual aids
 - evidence of rehearsal
- B. Quality of delivery**
 - rapport with and sensitivity to audience
 - use of visual aids
 - use of eye contact, voice, speed of delivery, gestures, movement
 - handling of questions
 - clarity and coordination of group delivery
- C. Language**
 - accurate and appropriate use of vocabulary, structures and register
 - concise and clear expression of ideas