Enhancing student learning, participation and accountability in undergraduate group projects through peer assessment

R. Subramanian*
e-mail: rajesh.subramanian@bothocollege.ac.bw

M. Lejk*
e-mail: lejk.mark@gmail.com
*Botho College
Gaborone, Botswana

Abstract
This article reports on a study undertaken into undergraduate student group projects at Botho College, Botswana. At the college, group projects are assessed entirely by the supervising tutor. This has led to some dissatisfaction on the part of students and staff when students’ contributions to the project are unequal and this is not recognised in the assessment. To overcome this, a trial was conducted in which a group of students performed peer assessment of the contributions of individual group members to the project. These assessments were compared to tutor assessments. The results are discussed together with feedback from students and staff and the implications for Botho College and higher education in general are examined.

Keywords: group projects, group assessment, peer assessment, self-assessment, accountability, co-operative learning

INTRODUCTION
Over the past 20 years group project work and its associated assessment has become one of the most common forms of student-centred activity. Whereas the benefits of working and learning in groups are widely accepted, the assessment is a much more difficult and controversial issue. Many methods of assessment have been reported with the vast majority using some form of peer assessment as a means of differentiating between individual student contributions (Baker 2008, 183–209; Lejk, Wyvill and Farrow 1996, 267–280). In some of the methods, an element of self-assessment is included in a similar manner to the peer assessment. Peer and self-assessment are practices which can be performed outside of the group environment where they are applied to work produced by individual students. There is an extensive literature on these latter areas, but the application of peer and self-assessment to group work is a more recent phenomenon. A comprehensive review of student involvement in assessment can be found in Falchikov (2005). This work provides a convincing rationale for the involvement of students in assessment and gives a comprehensive
overview of self, peer and collaborative assessment. Throughout, Falchikov stresses the need for guidelines and thorough preparation, without which ventures into this area will meet with limited success.

In this article the concept of peer assessment is examined and much of the discussion is based on a review of peer assessment by Topping (1998, 249–276). An attempt is made to classify peer assessment into different types. It is argued that peer assessment within groups is a very different issue to other types of peer assessment and the reasons for this difference are discussed.

The article also briefly examines self-assessment in a group setting and suggests that many of the considerations which apply to peer assessment in a group setting also apply to self-assessment.

Some relevant research about the practicalities of peer assessment within groups is discussed. This is followed by a report of an intervention at Botho College, Botswana, in which peer and tutor assessment of contributions to a group project are compared and reactions to the intervention by tutors and students are reported.

The reasons for assessment in groups are then analysed and the benefits of group assessment are discussed.

**PEER ASSESSMENT**

There is a substantial body of literature on peer assessment. The first point that must be made is that the term ‘peer assessment’ covers a diverse range of assessment experiences. A review by Topping (ibid., 249–276) concludes that peer assessment in higher education holds much promise especially in the area of formative assessment in the ‘cognitive, social, affective, transferable skill, and systemic domains’ but that much further research needs to be undertaken through properly planned interventions. Topping (ibid.) developed a useful 17-point typology of peer assessment in higher education, which reflects the parameters of variation between reported peer assessment experiences. These variables are listed below with the range of variation shown in brackets:

1. Curriculum area/subject (all)
2. Objectives (of staff and/or students? time saving or cognitive/affective gains?)
3. Focus (quantitative/summative or qualitative/formative or both?)
4. Product/output (tests/marks/grades or writing or oral presentations or other skilled behaviours?)
5. Relation to staff assessment (substitutional or supplementary?)
6. Official weight (contributing to assessees final official grade or not?)
7. Directionality (one-way, reciprocal, mutual?)
8. Privacy (anonymous/confidential/public?)
9. Contact (distance or face to face?)
10. Year (same or cross year of study?)
11. Ability (same or cross ability?)
12. Constellation of assessors (individuals or pairs or groups?)
13. Constellation of assessed (individuals or pairs or groups?)
14. Place (in/out of class?)
15. Time (class time/free time/informally?)
16. Requirement (compulsory or voluntary for assessors/ees?)
17. Reward (course credit or other incentives or reinforcement for participation?)

Two of these parameters (12 and 13) are the constellation of assessors and the constellation of assessees and these aspects allow for a general classification into the following four types:

- **Type 1:** Peer assessment of work produced by an individual by another individual. This occurs frequently in the assessment of writing of which there is a voluminous amount of literature (e.g. Pain and Mowl 1996, 19–31).

- **Type 2:** Peer assessment of work produced by an individual by a collection of other individuals. This usually involves the assessment of presentations or posters and measures of validity use a correlation between a tutor mark and the average of the marks awarded by the peer group (Saito and Fujita 2009, 149–171).

- **Type 3:** Peer assessment of work produced by a group of students by a collection of other individuals (sometimes a collection of other groups) (Freeman 1999, 289–299). As with Type 2, comparisons of tutor grade with average peer grade have been made.

- **Type 4:** Peer assessment of the contribution of an individual to a group assignment by the other members of the same group. This is in general a more recent practice. Although Fineman (1981, 82–93) did report such an exercise most reports appeared from the 1990s onwards and the volume of literature is accelerating (e.g. Scott, Van der Merwe and Smith 2005, 61–70). By its nature, this type of peer assessment is more difficult to analyse for validity and reliability than the other forms. Topping (1998, 260) concludes that ‘the outcome data on this type of peer assessment are limited to student perceptions’. Falchikov (2005, 200–221) dedicates a separate chapter to this Type 4 classification in order to distinguish it from Types 1, 2 and 3.

The results of studies into Types 1 to 3 are rather varied. Without really looking at the literature, it could be predicted that peer assessment of a multi-choice test using the correct answer scheme would be more valid and reliable than peer assessment of an open-ended essay where peers were required to simply use their judgement in arriving at a grade. Most assessments lie between these extremes and some general observations can be made.

**Formative benefits**

The overwhelming benefit of Types 1 to 3 peer assessment is formative. The process
develops high-level critical skills and involves the students in the learning process through assessment. This comes through from all the reported studies (see Falchikov 2005 for a review).

**Nature of assessment criteria**

The validity of the peer assessment is influenced by the nature of the assessment criteria. For example, MacAlpine (1999, 15–25) found that assessment of student presentations was closer to tutor assessments when the criteria involved four separate categories compared to an overall holistic impression. However, Kwan and Leung (1996, 205–214) obtained low agreements between peer and tutor scores when they used six categories which included a total of 35 separate indicators and a six-point grading scale.

**Development of peer assessment skills**

Students should be introduced to peer assessment over a period of time in absorbable and achievable steps and preferably with some input into the development of the assessment criteria (Langan, Wheater and Shaw 2005, 21–34). The ability to assess peers is itself a learning process. This suggests that an eighteenth variable should be added to Topping’s topology: level of assessor training.

**Qualitative assessment**

Studies of reliability and validity have concentrated on quantitative assessment. There is a general tendency for a narrower spread of marks in peer assessment than in tutor assessment. Studies of the reliability and validity of qualitative, formative assessment are needed.

**Substitution for tutor assessment**

Because of the variation in reports, and the potential sensitivity to even slight variations in the elements of Topping’s (1998) 17-point typology, the use of peer assessment to replace summative tutor assessment to any great extent remains a controversial practice.

**PEER ASSESSMENT OF THE CONTRIBUTION OF AN INDIVIDUAL TO A GROUP ASSIGNMENT BY THE OTHER MEMBERS OF THE SAME GROUP (TYPE 4)**

The preceding observations apply to peer assessment of Types 1, 2 and 3. Type 4, peer assessment of the contribution of an individual to a group assignment by the other members of the same group, has been omitted. In the current article it is argued that Type 4 peer assessment is so very different to the other types of peer assessment as to warrant totally separate consideration. Whereas similar measures for validity and reliability can be devised for Types 1 to 3, the same measures are either inappropriate or impractical for Type 4 and the environment in which the assessment is carried out.
can give the assessors a totally different ‘agenda’ through which they carry out the assessment.

Topping’s review is about the complete range of peer assessment activity in higher education. As such, Topping’s analysis of peer assessment in group projects (Type 4) is rather brief. There are several considerations which apply to Type 4 peer assessment, some of which force it to be treated very differently from Types 1 to 3 and others which can be applied to Types 1, 2 and 3 with a similar rationale. These considerations are discussed below.

Assessment of product or process

Types 1 to 3 invariably involve the assessment of some product by peers; for example an essay, a test, a presentation or a poster. In the literature on Type 4 peer assessment, the assessment of the product produced by the group is always assessed by a tutor or expert. The members of the group assess each other on individual contributions to that group product, that is they assess a process.

Tests of validity

The way this process is assessed varies considerably (Baker 2008, 183–209; Lejk et al. 1996, 267–280). Some methods use a holistic approach, whereas others use a category-based approach. In the latter approach the categories themselves can vary considerably. For example, Falchikov (1998, 327–339) sub-divides the categories into task functions and group maintenance functions, many of which can be linked to interpersonal traits. Conway, Kember, Sivan and Wu. (1993, 45–56) use task-related categories (e.g. literature search, writing the report, etc.) while others use combinations and/or variations of these (Scott et al. 2005, 61–70). In all these approaches the assessment of an individual’s contribution to the process is rarely undertaken by a tutor or an expert. There have been some studies of assessment of contribution to relatively short group activities by trained assessors and experts which warn respectively against self-assessment and non-professional assessment of specific interpersonal traits (Shechtmann 1992, 31–39). However, in extended group projects ‘it is impossible for a tutor to have the necessary intimate knowledge of each individual student’s contribution to a project’ (Lopez-Real and Chan 1991, 67–79). It is therefore impossible to have an objective test of validity. Studies have been limited to measuring the extent of agreement between members of the same group which is really a test of reliability.

‘Zero-sum’ games

One common way in which peer assessment is used is to apply it to the mark obtained for the group product. Most of the variations on this theme are examples of a ‘zero-sum’ game in which an increase in marks for one individual must be accompanied by a decrease in marks for another.

A common way of doing this is by distributing a pool of marks (Lejk et al. 1996, 267–280), a method which has been adapted to yield many variations. An example
follows:
In a group of four students, the group product is awarded a mark of 60 per cent by the tutor. The group are given 100 points to divide up to reflect individual contributions, resulting in Student 1 being awarded 30 points, Student 2 being awarded 20 points, Student 3 being awarded 25 points and Student 4 being awarded 25 points.

Applying these weightings to the 60% group score:
Student 1 gets 30/25 or 1.2 x 60 = 72%
Student 2 gets 20/25 or 0.8 x 60 = 48%
Student 3 gets 25/25 or 1.0 x 60 = 60%
Student 4 gets 25/25 or 1.0 x 60 = 60%

It can be seen that applying peer assessment in this way gives a spread of marks and prevents students benefiting unfairly by giving each other all high scores. There are many variations on this theme but most adopt the same basic principle of the ‘zero-sum’ game.

Other problems of validity
The relatively few reports in which the peer assessment of process does not influence the assessment of product (i.e. it is not equivalent to a ‘zero-sum’ game) still suffer from the same problems of validity already mentioned. In these reports all the members of the group receive the same grade for the product which eliminates the problems of the ‘zero-sum’ game and does therefore seem to be going in the right direction. However, other problems with peer assessment have been raised such as cultural climate (Lopez-Real and Chan 1991, 67–79) gender issues and friendship groups (Langan et al. 2005, 21–34). These considerations will, however, also be relevant to Types 1, 2 and 3 peer assessment.

The wider assessment picture
When students undertake assessments in groups, the wider assessment picture needs serious consideration. Students can schedule the completion of a set of individual assignments to fit in with their life outside university. With group assignments, it is more complicated. Students can be involved in group assignments in different modules at the same time. In addition, students are sometimes working with different sets of peers in the different assignments. This becomes a strategic nightmare for students when it comes to organising meetings outside class.

Timing of peer assessment
In Types 1, 2 and 3 peer assessment, the assessment is generally performed after the product has been completed although one can imagine situations in which it could be applied to draft intermediate products. In groups this consideration takes on great significance. If peer assessment of individual contributions is only done at
the end of the group exercise (as is often the case in the reported studies, e.g. Lejk et al. 1996, 267–280), then an obstructive or errant group member may have already seriously damaged the performance and/or atmosphere of the group. It makes sense to have ongoing, possibly informal, peer assessment throughout the course of the group exercise so that any problems are highlighted quickly and dealt with. In this case, peer assessment is being used as a control mechanism to monitor the process.

**Manipulation and abuse**

Some methods of peer assessment in groups can be manipulated or abused by students, especially those which are variations of the ‘zero-sum’ game. For example, it can sometimes be to a group’s advantage to have a very weak member. The other group members could conceivably do the bulk of the work to a high standard and the weak member not be allowed to do much at all. This weak member would presumably suffer in the peer assessment and the other students would gain by this student’s loss. Not a very democratic scenario but a distinct possibility. It argues the case for close monitoring of groups by the supervising tutor. Naturally, individual assessments are open to various forms of manipulation or abuse, but in group assessments they are often extremely difficult or impossible to detect.

**Method of group formation**

The method of group formation can influence the performance of the group. Groups can be formed by self-selection (which usually leads to friendship groups), by random selection or by some other form of engineering, for example by personality type or ability. Lejk et al. (1999, 5–14) showed that high ability students obtained considerably lower grades when placed in mixed ability groups than in streamed groups whereas the reverse is the case for students at the lower end of the ability range. This study did not include peer assessment; the grade awarded to the student was the same as the grade awarded to the group. However, it does show that careful consideration needs to be applied to the relative ability of the group members. Studies of group dynamics show that certain personality types work better together in groups than others. For example, Belbin (1981; Aritzeta 2007, 96–118) has studied management teams and derived a number of group working characteristics, different combinations of which lead to differences in group cohesion.

**Length of time allocated**

The length of time allocated to a group assignment is important. If the groups are formed on a basis other than that of friendship, then time has to be allocated for the group members to get to know each other. There are a number of theories about the phases of group development but their basic content is quite similar. A typical catchy example is provided by Tuckman and Jensen (1977, 419–427) where the five stages are labelled ‘forming’, ‘storming’, ‘norming’, ‘performing’ and ‘adjourning’.

Clearly each of the first three stages requires time *in addition* to time allocated for performing the tasks. In a short group assignment, the students may well feel under
so much pressure that the ‘storming’ and ‘norming’ phases are simply missed out and the group members hurriedly perform individual bits which are then hurriedly pieced together to give a product which lacks integrity.

**Size of group**

The size of the group is important. Three, four and five member groups are commonly found; six is not an infrequent number; and large groups of 10 to 12 and up to 15 members have been reported. The consideration of group size is related to the previous ones of the length of time allocated to the group exercise, and the method of group formation. The larger the group, the longer will be the period of time needed for the group to gel, especially if the members are mostly unfamiliar with each other at the beginning. Peer assessment is more difficult in a large group as there is a tendency to split into subgroups and it is quite possible for one subgroup to have little idea of the quality and contribution of the work of individual members of the other subgroups.

**Differences in aspirations between group members**

This is an obvious problem with group work especially if groups are engineered in some way. The situation may well arise where one or more of the students in a group are aiming to do really well, for example aiming for a first-class degree, whereas others have more modest aims, perhaps being happy with a lower second. Alternatively, the members of a group may have other assessments at the same time, some members having more than others with assessments carrying differing contributions to final module grades. Clearly, this may lead to varying levels of importance attached to the group work among the group members.

**Confusion between effort and quality of work**

In traditional assessment of individual work, the effort expended by the student is largely ignored. It is the quality of the final product which is assessed. Within group work, the quality of the final product is normally assessed by a tutor. The peer assessment element is performed before the tutor assessment and there is often confusion as to whether the students are assessing each other on the effort expended or the quality of the work produced. The dilemma is easily seen as is the difficulty in distinguishing between the two aspects from the students’ point of view.

**THE RELATIONSHIP BETWEEN GROUP ASSESSMENT AND PEER ASSESSMENT**

The previous analysis highlights some missing variables in Topping’s (1998) typology when applied to Type 4 peer assessment. Such variables include size of group, method of group formation and length of assessment. Types 1, 2 and 3 peer assessment are themselves complex phenomena, but Type 4 is even more so because of the interaction and interdependence of the assessors during the completion of the
assessed task(s).

There are two concerns here. Firstly, conclusions reached about Types 1, 2 and 3 peer assessment may be carried across erroneously to Type 4. Careful researchers and practitioners would not fall into this trap, but others might. The message has to be stressed that peer assessment within groups is very different from other types of peer assessment. Secondly, group assessment may become inextricably linked with peer assessment. In other words, an impression may be given that group assessment must involve some form of peer assessment. Again, it must be stressed that group assessment is a complex business which may involve peer assessment but equally may not.

It should be noted that the concerns described are not intended as criticisms of Topping’s typology, which is a clever and useful way of classifying peer assessment experiences. There is an overlap between research into group assessment and research into peer assessment and the typology has provided a much-needed framework for analysis of this overlap.

**SELF-ASSESSMENT**

Self-assessment of work performed by individuals has a longer history than equivalent peer assessment. A major review and meta-analysis have been undertaken (Boud and Falchikov 1989, 529–549; Falchikov and Boud 1989, 395–430). Self-assessment of contributions to group work has always been reported as an adjunct to the associated peer assessment and useful analysis of the self-assessment aspect is infrequent. Falchikov (1988, 327–339) found that agreement between self ratings and peer ratings is considerably less than agreement between peer ratings, and other studies indicate that there is a tendency for stronger students to underrate themselves and weaker students to overrate themselves in comparison with assessment by their peers (Lejk and Wyvill 2001b, 551–561; Violato and Lockyer 2006, 235–244). A number of reports of group assessment do not include a self-assessment element, possibly for these reasons (Scott et al. 2005, 61–70).

Some of the factors which have been presented in the article as having an influence on peer assessment within a group setting are also relevant to self-assessment. The arguments about tests of validity, ‘zero-sum’ games, category-based assessment, manipulation and abuse, method of group formation, differences in aspirations between group members and confusion between effort and quality of work apply equally well to self-assessment and peer assessment within a group setting.

**SOME REPORTED INTERVENTIONS ABOUT PEER ASSESSMENT WITHIN GROUPS**

In a series of reported interventions at the University of Sunderland, United Kingdom, Lejk and Wyvill (2001a/b) addressed a number of research questions. Three of these questions follow together with the major findings of the research.
Question 1

*When peer assessment takes place within a group assignment, what differences occur in the final distribution of marks when the peer assessment is performed holistically and when it is broken down into categories?* (Lejk and Wyvill 2001a, 61–72)

The main findings of this intervention are presented below:

- Both methods of peer assessment increase the correlation with individual grades obtained from other individual assignments, as compared to awarding all the group members the same grade.
- Holistic assessment gives a higher proportion of equal distribution of marks.
- Holistic assessment gives a higher proportion at the extremes. In other words, students who perform particularly well or particularly poorly are highlighted to a greater extent in holistic assessment.
- Students show a tendency to perform category-based assessment in a holistic way, frequently awarding the same mark for all the categories.

Question 2

*When self-assessment is included with peer assessment in a group assignment, how does the self-assessment of an individual compare with the peer assessment of that same individual?* (Lejk and Wyvill 2001b, 551–561)

The main findings of this intervention follow:

- Better students tend to under-assess themselves.
- Weaker students tend to over-assess themselves.
- Removal of self-assessment increases the spread of marks.

Question 3

*When peer assessment is used within a group assignment, what differences occur when the peer assessment is performed secretly and when it is agreed as a group?* (ibid.)

The main findings of this intervention are presented below:

- Agreement between peers is less when done secretly.
- The spread of marks is higher when done secretly.

These results have obvious implications for the conduct of peer assessment within groups. Students are more discriminating when performing peer assessment in secret and the inclusion of self-assessment seems to skew results ‘unfairly’. The debate about holistic and category-based peer assessment is not so clear-cut. Category-
based peer assessment is obviously more useful for feedback to students whereas holistic peer assessment seems to support the objectives of team work better than category-based peer assessment.

**GROUP PROJECTS AT BOTHO COLLEGE**

Group projects are common at Botho College. The courses are generally in computing and, as computer systems are designed and built by teams, it is good practice to prepare students for working life by exposing them to group project work. The method of assessment is standardised for all group projects. The students work together on a project and the final assessment is performed by the tutor. Ten percent of the marks are awarded for group co-ordination and the tutor awards these marks by observing the group whenever the opportunity arises and by questioning in a group presentation after the project has been submitted. However, some students have commented that this has led to an unfair allocation of marks where some group members can ‘hoodwink’ the tutor into awarding more marks than they deserve, whereas others who contributed a great deal do not seem to be rewarded for their efforts. It was therefore decided to conduct an intervention in which students would be allowed to peer assess the other members of the group and award marks in the same way as the tutor.

The methodology used was a mixture of quantitative analysis of results and analysis of student and staff attitudes to the intervention using questionnaires. The group co-ordination mark (worth 10%) which is normally awarded by the tutor was also awarded by the students using peer assessment. As the tutor awarded this 10 percent of marks by using categories, the same method was used by the students. The categories were: distribution of group tasks; communication with members; listening to other points of view; attitude, attendance and readiness; and finally effectiveness. Students awarded each other marks out of 10 for each of these categories. The students did not assess themselves and the assessments were performed in secret before being handed to the tutor. The tutor averaged the marks for the individuals and converted the score into a mark out of 10.

In addition the tutor awarded marks for group co-ordination in the same manner. It was the tutor mark which counted for the final assessment.

An attitude survey was conducted to gauge the extent to which both students and staff regarded the two methods in terms of fairness. Both staff and students were asked to choose one of the following five options for each method:

- whether they found the method fair;
- whether they found the method fair but with a few concerns;
- whether they found the method unfair;
- whether they were undecided;
- whether they were confused.
Results were analysed for 253 students in groups of four to six and 14 tutors. The results were as follows:

- the mean tutor mark was 6.14 out of 10, standard deviation 1.33;
- the mean peer assessed mark was 7.24 out of 10, standard deviation 1.38.

Hence the peer assessed mark was higher by a factor of 11.1 per cent. Using a two-sample Z-test the two means are significantly different for an alpha value of 0.05 (Z-value 9.13, 2-tailed p-value < 0.0001):

- 62% of students agreed that the tutor assessment is fair or fair with a few concerns;
- 89% of students agreed that the peer assessment is fair or fair with a few concerns;
- 43% of tutors agreed that the tutor assessment is fair or fair with a few concerns;
- 86% of tutors agreed that the peer assessment is fair or fair with a few concerns;
- 2 of the 238 students were undecided and 14 were confused by both methods. Indecision and confusion were not displayed by any of the tutors. Hence more tutors and students agreed that the peer assessment method is fairer than the tutor assessment method.

**DISCUSSION**

These results highlight both the advantages and the pitfalls inherent in peer assessment within groups. Firstly, allowing free reign in peer assessment can be dangerous. If students see that the marks which they award will be unchallenged and unadjusted, then they will have a tendency to award high marks. They become strategic in their assessment and this observation is apparent in this intervention where the marks awarded by peers were significantly higher than the marks awarded by tutors, even though the students knew that it was the tutor mark which would be awarded. This is where the application of a ‘zero-sum’ game principle is useful, in which awarding a high mark to one peer immediately lowers the final marks awarded to other peers. Future interventions are planned at Botho College in which a ‘zero-sum’ mechanism will be applied. However, both students and staff seem to regard the peer assessment method as fairer than the tutor assessment. It is striking to see that the tutors regard their own assessments of group co-ordination in a more negative light than their students do. This could be explained by the fact that they are not present during the bulk of the activities involved and so do not feel in a position to judge student contributions fairly. However, they are very supportive of the peer assessment mechanism where the students themselves do the judging. Attitude surveys from other studies also support this finding (Lejk and Wyvill 2002, 569–577).

It is intended to analyse the data from the current study further. It is hoped to examine gender and age differences in peer assessment and also variations in agreements between tutors in peer and tutor assessment.
Given the many factors involved and the potentially problematic nature of group assessment, it is tempting to regard the whole thing as too complicated to contemplate and revert to a complete diet of individual assessment. Even champions of group project work reveal a tension with the assessment. For example, Falchikov and Magin (1997, 385–396) reporting on two group-based case studies comment that ‘in both cases, the marks awarded by students made a contribution to their final grades for the project, though not so large a contribution as to undermine cooperative learning or distort marking’.

There is a danger that assessed group activities may be given to students for the wrong reasons. With the increasing numbers of students in higher education worldwide in recent years, the pressure on tutors has increased considerably, especially in the area of assessment. The replacement of individual assessments with group assessments is usually accompanied by a decreased marking load. This, on its own, is not a good reason and can lead to all sorts of problems, not least an excessive amount of group assessment in a student’s diet. However, there are good reasons for doing group work.

Cooperative and collaborative learning are good ways to learn about the subject matter. There are numerous studies which support this view. It is also an ideal way of developing interpersonal skills. It is a good way of exposing students to large projects, similar to ones found in the real world, which are undertaken in teams. This is particularly relevant to vocational courses. For example, in the computing world, the products are nowadays so large and complex that their design and development necessitate a team approach. It is therefore a sound idea to give students some exposure to this sort of activity. Some examples of ‘real-world’ group projects add to the realism by involving local employers (Goldfinch et al. 2001, 367–380). Although such group work need not be assessed, it is difficult to imagine how students would take these activities as seriously as their other assessments which do contribute to their awards.

The assessment of such group work, although problematic, also has considerable benefits. When students undertake Type 4 peer assessment, they feel that they are an important part of the assessment process. They take on more responsibility and motivation improves. They become better prepared for independent learning. They develop evaluation skills which are useful for future employment and life in general. They treat assessment as an integral part of the learning process. In addition, if peer assessment is included, it would encourage accountability. From an institutional point of view, there are resourcing benefits. If conducted properly peer assessment can replace some tutor assessment.

**CONCLUSION**

There are many factors which must be considered in peer assessment of individual contributions to group assignments which do not apply to other types of peer assessment. As a result, the practice must be deliberately distinguished and examined
Enhancing student learning, participation and accountability in undergraduate group projects

separately. When peer assessment is performed in a group environment, considerable care must be taken in the management of the assessment process. The study reported in this article indicates that peer assessment of group processes in which the results stand in their own right without any adjustment is a dubious process. However, both staff and students involved in the study regarded peer assessment of this aspect as fairer than tutor assessment on its own. Further studies are needed to investigate the effect of a ‘zero-sum’ game mechanism on peer assessments in group projects at Botho College. It is intended to analyse the data in this study further to examine gender and age differences in peer assessment and variations in agreements between tutors in peer and tutor assessment.

REFERENCES


grades from group assessments. *Assessment & Evaluation in Higher Education* 21(3):
267–280.

——. 1997. Group learning and group assessment on undergraduate computing courses in
higher education in the UK: Results of a survey. *Assessment & Evaluation in Higher

——. 1999. Group assessment in systems analysis and design: A comparison of the
performance of streamed and mixed-ability groups. *Assessment & Evaluation in

Lejk, M. and M. Wyvill. 2001a. Peer assessment of contributions to a group project: A
comparison of holistic and category-based approaches. *Assessment & Evaluation in

——. 2001b. The effect of the inclusion of self-assessment with peer assessment of
contributions to a group project: A quantitative study of secret and agreed assessments.

——. 2002. Peer assessment of contributions to a group project: Student attitudes to holistic
569–577.

mathematics education course. *Assessment & Evaluation in Higher Education* 24(1):
67–79.


Saito, H. and T. Fujita. 2009. Peer-assessing peers’ contribution to EFL group
149–171.


Shechtmann, Z. 1992. Interrater reliability of a single group assessment procedure
administered in several educational settings. *Journal of Personnel Evaluation in


Tuckman, B. W. and M. A. Jensen. 1977. Stages of small group development. *Group and

Violato, C. and J. Lockyer. 2006. Self and peer assessment of pediatricians, psychiatrists
and medicine specialists: Implications for self-directed learning. *Advances in Health
Sciences Education* 11(3): 235–244.