

Evaluating short-term and long-term peer assessment of Student Teamwork

Michael Sergi

Division of Communication and Education

University of Canberra

Canberra, ACT 2601 Australia

Email: Michael.Sergi@canberra.edu.au

Abstract

The lack of individual reward for individual effort is a major concern for many university film and video students undertaking group-based projects. Peer assessment is often used to derive individual marks from group projects, and because it goes some way towards mimicking professional practice. However, if there is only one group project that is part of a subject's assessable work, any mismatch of students, in terms of skills, commitment and personality, can result in some students receiving an undeservedly harsh assessment from their peers, which can affect their final grade. Long-term peer assessment, where students undertake several small group projects in a semester, each with a different set of students, has been shown to be a useful instrument for deriving a more moderated peer assessment mark, not least because it acts to counteract the potential pitfalls of mismatched students.

Key words: *Peer assessment, Team work, Group projects*

Introduction

Small group interactions provide opportunities for every group member to be successful. Members are not competing with one another but are working together to achieve a common goal. (Brown 1996). The benefits of having students work together in teams ...[include]...the development of teamwork skills, enhanced student interest and motivation, the potential for students to learn from each other, the opportunity for multicultural experiences, and the opportunity for a student to explore more deeply an aspect of a project of great personal interest (Bacon, Stewart and Stewart-Belle 1998; Dommeyer 1986; Williams, Beard and Rymer 1991).

Film and television production is intrinsically collaborative by nature. Production teams of various sizes, sometimes numbering in the hundreds, make the majority of film and television products created around the world, although very small teams or even individuals have made some outstanding films. Within the professional industry, being an efficient and effective team member is essential for career success. Consequently, it is a core responsibility of university film and television production courses not only to teach the theoretical and technical aspects of filmmaking, but also to teach students how to improve their management of the filmmaking process; how to work effectively and efficiently in teams/groups; and how to assess self and peers alike. At the professional level, the work of individuals is commonly peer assessed. The Academy Awards are just another form of peer assessment: 'the situation of the professions...[is]...that most assessment is by peers' (Boud and Tyree, 1995).

Research by Holpp and Phillips (1995), and Chapman and van Auken (2001), although conducted in business courses, indicates that students instinctively tend to focus on the product rather than the process, "A common problem for students is that they often overlook the importance of process and instead focus on tasks and outcomes. This is particularly the situation when students work on group projects", (Chapman and van Auken 2001). My own experience with film and television students supports their research. Assessing the process would indicate to students that teachers value the process as much as the end product. It is important to instil in students an appreciation for process, because, sadly, students only seem to appreciate those elements that are assessed. Beatty, Haas, Sciglimpaglia (1996) found that if a group-based project carried a low assessment value, it was of less importance to students, and impacted negatively on their individual contributions.

Small group film and television production projects, where group member numbers average somewhere between 4 and 7, are a common way of teaching the production aspects of film and television. However, the assessment of the individual's performance within a group, as apart from the project itself, has long been a contentious issue: "Group performances are inherently difficult to grade, with a major problem being the assignment of grades to individual group members" (Beatty, Haas, Sciglimpaglia, 1996). "Previous research suggests that perceptions of work inequity translate into perceptions of grade inequity and to overall perceptions of unfairness in grading" (Beatty, Haas, Sciglimpaglia, 1996; Forman and Katsky 1986; Williams, Beard and Rymer 1991). "Students have expressed fear that their individual grades will be compromised or that the work will not be distributed fairly among members of the group (e.g., Chapman and van Auken 2001; Comer 1995).

Teachers and students understand that how well students work together as a group, their blending of talents and management of personality differences will have a significant bearing on the quality of the project (Beatty, Haas, and Sciglimpaglia 1996). Although they were studying marketing students, Bacon, Stewart and Stewart-Belle (1998) found that the average ability of each team member had a noticeable effect on the team performance and the grade achieved for the group project. Chapman and van Auken (2001) found that students with high Grade Point Averages were more anxious about working in groups that consisted of students with lower GPAs. For Lembke and Wilson

(1998) "Highly productive teamwork requires that team members recognise the team as a unit and as an attractive work arrangement" (p.927).

However, teacher and student apprehension regarding the assessment of group projects is common. Research has shown that students cite the lack of individual reward for individual effort as a major concern when undertaking group-based projects. "Instead, their [the students they surveyed] experiences confirm Slavin's (1989) assertion that group rewards based on a single group product may set up conditions where one or two members do most of the work" (Colbeck, Campbell and Bjorklund, 2000) And as Grant (1994) observed: "many academic staff are hostile to group projects because of the assessment implications, and students' criticisms are mainly of the assessment procedures, as many students feel that their individual efforts are not sufficiently rewarded". This is particularly the case when a single group project is the only group project in a semester, or even worse, when the major assessment component of a subject is a group project.

Brown (1996) also supports this position: "A basic complaint surfaced: students' contributions to the project were not equal, yet each group member received the same grade. Some students felt this was not fair. I tended to agree and felt I was not adequately assessing the students' individual efforts on the projects". Hard working competent students have every right to object when others, who contribute little to the project, or worse, disrupted their efforts, receive the same marks. Students are aware that the quality of each group member's contribution has a significant effect on the quality of the project, and therefore, should be included in the assessment. The essence of the students' complaint was that only the completed project and not the process was assessed. Furthermore, when it comes to film and television production, managing the process is a significant part of the project, and should be a key element in the students' learning. Not only should students be learning how to operate technical equipment, and direct actors (if they are making a drama), but they should also be learning how to function as a professional film and television crew, which involves adopting professional production practices. For many students, however, their understanding of the process aspects of group work is typically confined to its relationship with the project mark.

Students quite rightly fear that if the process is assessed it might not be done fairly. Chapman and van Auken (2001) found that peer evaluations, and other methods of assessing individual contribution to the group project were not frequently used. Teachers are also apprehensive about assessing the process because of the increased workload and because existing evaluation methods, such as peer assessment, might be perceived as being unreliable. Chapman and van Auken (2001) realised that while peer assessment is a good instrument for assessing process it "will never completely remedy the inequity problem". When discussing the self-assessment component of some peer assessment methods Boud & Tyree (1995) found that, "students tend to rate themselves more favourably than they are rated by their peers".

However, because the majority of students' project work is completed outside class time, which makes it unsupervised, a teacher has no way of assessing an individual student's participation and contribution to a project without the information provided by peer assessment. Only the students themselves are positioned to know about each member's contribution. "As elusive as it may be, insight into a group's interrelationships is necessary if the instructor is to achieve equitable grading" (Beatty, Haas and Sciglimpaglia 1996).

Both teachers and students know that a group member who is lazy, or disruptive, has poor ability, or is overly ambitious can have a negative impact on the quality of a group's work. They also know that some groups, for various reasons, do not function well, because group work is stressful and conflict within the group is common (Lerner, 1995; Brown, 1996). Consequently, if an ambitious student is unlucky enough to be in a

dysfunctional group, where their desire to do well is resented, they might not only receive a poor project mark, but also a negative peer assessment that might not be deserved. This is one of the major short-comings of having a single non-peer-assessed group project in a semester.

Some Existing Methods Of Peer Assessment

Lejk et al. (1996) presents a good summary of some current assessment methods used to assess group work. These methods can be divided into two categories: those that distribute a single project mark equally among all group members and do not assess the process at all; and, those that assess the project and the process either by using a form of peer assessment, or some other method (Falchikov 1991). Of these two broad categories, the most commonly used is where the group mark is distributed equally among all group members, with no assessment of individual participation and contribution (Mello, 1993). Students are most hostile towards this method. Typical complaints are: "For at least two team leaders, assuming the leadership role meant doing most of the project work. One told us, 'Mostly I find it gets done better if I do it myself.' Another completed a friend's task rather than jeopardise the friendship." (Colbeck, Campbell and Bjorklund, 2000).

There are several variations of the above two methods that are worth exploring. One is where the teacher designs a group project made up of several individual components and students submit their work independently. An example of this would be the script breakdown, scheduling and budgeting of a film. Although, the outcome of one, the budget, is dependent on the outcome of the others, the script breakdown and schedule, there is a large enough degree of autonomy between them to assess them individually. In such a case, the amount of genuine group work is limited. In essence, this is not a group project at all, but a study group. The amount of teamwork skills acquired by students doing such a project would be minimal.

In a variation of this method, an aspect of genuine group work that carries a low mark, e.g. producing a storyboard, is added for which the students receive a group mark that is distributed equally, while the separate components (the script breakdown, scheduling and budgeting) are assessed individually, (Hindle, 1993). The total mark a student receives is the sum of the group mark and their individual mark. In this method, the teacher must ensure that the individual tasks are equally complex and demanding, and this may not always be possible. Any perceived inequity in the complexity of the individual components, either justly or unjustly, can cause friction between group members.

Another group assessment method is the 'Distribution of a Pool of Marks' (Habershaw et al., 1993). Here the group receives a mark for the project, which is then multiplied by the number of group members, e.g. 55 (project mark) x 4 (number of group members) = 220 (pool of marks). This pool of marks is then given to the group to distribute amongst themselves, usually in an open group situation. Alternatively, individual group members might distribute the marks in private, and the teacher then tallies the results and averages them out (Lejk et al., 1996). This method may cause friction between group members if they are allocating marks in an open group situation, because some students might feel that they deserve a higher mark than the other members are prepared to give them. Or students may decide to split the marks equally because of peer pressure, thus rendering the exercise futile. If the students distribute the marks in private, and the teacher tallies the results and averages them out, it is still possible for a student to deduce that their fellow group members gave them a low mark, which they might believe is unfair. This may also cause friction amongst the cohort.

More sophisticated methods of group assessment seek to determine the actual level of individual participation and contribution. One such method is where the teacher uses individual reports written by group members, and/or vivas (oral exams), to determine

the level of individual participation and contribution (Earl, 1986). This evaluation is then considered against the project mark to determine a final mark for each student. The teacher must exercise careful judgement because contradictions between reports often arise. There is a considerable increase in teacher workload with this method, as the teacher must read all the individual reports, or give vivas to each student in private, and then determine a mark for each student from the information they have gleaned.

In the Habershaw (et al., 1993) method, the teacher adjusts a group member's mark depending on the peer assessment of that group member. The sum of the project mark and the peer assessment mark cannot be greater than the project mark. A shortfall of this method is that a talented student, in a group of less able students, can never achieve a mark higher than the group project mark. Since the quality of the project relies on the ability of the group as a whole, the project mark is unlikely to be as high as if the talented student was working with students of similar ability. Also even if the students of this group valued the contribution of the talented student they could not give them a higher mark for participation and contribution, because the group project mark is the upper limit.

In the Goldfinch method, each group member's mark is made up of a percentage of the project mark, plus a peer assessment mark. This method allocates each student a peer assessment factor, which is derived by dividing a student's peer assessment mark by the group's average peer assessment. The Goldfinch method includes self-assessment as part of the peer assessment process. An algebraic formula is then applied to the project mark and the peer assessment mark to derive each group member's mark, (Goldfinch and Raeside, 1990; Conway et al., 1993; Goldfinch, 1994). Student's final mark = project mark X peer assessment factor.

With this method, a talented student working in a group of average students can achieve a final mark higher than the project mark, and higher than fellow group members, because peer assessment enables group members to reward an individual who performed above the group average. This form of peer assessment also allows group members to mark down a student who did not contribute their fair share. This method truly attempts to allocate a judicial mark for group participation and contribution.

However, because the Goldfinch method derives the student's final mark by multiplying the project mark by the student's peer assessment factor, it implies that there is a direct correlation between the student's peer assessment and the quality of the project. This is where a weakness in the methodology might exist. A talented student, whose drive for success could have had a significant influence on the quality of the project, but in doing so might annoy some group members, consequently they receive a poor peer assessment. In such a case, the talented student might receive an overall mark that is lower than the project mark, while the students who gave them a poor peer assessment reap the rewards of the talented student's hard work: "analyses clearly shows that, based on the perspectives of those most directly involved in the project, group members performances were not equal throughout the project" (Beatty, Haas and Sciglimpaglia 1996).

Furthermore the Goldfinch method places more emphasis on the student's individual contribution than it does on their group work skills. This is a subtle difference that becomes significant in student film and television production projects where students might swap roles during the project, e.g. they might operate the camera one day and record sound the next. This is not that uncommon, particularly in early projects when students are still trying to discover which task interests them the most. Because swapping tasks occurs, it is problematic attempting to discern, with any certainty, what each individual contributed and to what extent – did Sally really do all the camera work? Thus placing more emphasis on individual contribution than on group participation, as the Goldfinch method does, makes it less suitable for film and television production.

The Falchikov (1991) method separates the process mark from the project mark and students are allocated both. With the Falchikov method, the teacher assesses the project, while the process, both individual contribution and group participation, is peer assessed using a survey. This method differs from the Goldfinch method in that it allocates the process mark, which is individual, independently from the project mark, which is given to each member of the group. Falchikov divides the assessment of the process into two categories—(1) task functions and (2) group maintenance functions.

Of all the methods described, the Falchikov method is the most sophisticated, and most suitable to film and television production, because it recognises that the project and the process are separate yet interdependent components of group work. Marking only one element of group work, usually the project assesses only part of a student's work (Colbeck, Campbell and Bjorklund, 2000). Chapman and van Auken (2001) found that teachers infrequently attempted to assess individual group performance. To obtain a more accurate assessment of student work, both the project and the process should be assessed. Beatty, Haas and Sciglimpaglia (1996) resolved that, "both instructors and students desire realistic evaluation, fair grading, and equitable assignments of grades to group members".

While many of the methods outlined have considerable merit, especially the Falchikov and Goldfinch methods, their weakness is that they are all based on short-term peer assessment. Short-term peer assessment is where students complete only one project, and are members of only one group, in a semester. In situations where only one of the assignments for any given semester is a group project, a student might be disadvantaged if they are unfortunate enough to be a member of a dysfunctional group. Bacon, Stewart and Stewart-Belle (1998) found that the quality of team performance was predicated on the average ability of team members.

The above methods also fail to account for the potential generosity that is likely to occur when friends peer assess each other (Goldfinch, & Raeside 1990). There is also the possibility that personality clashes, or victimisation, might result in a student receiving an undeservedly harsh peer assessment.

Long-term peer assessment is where a student completes more than one project, and is the member of more than one student group, during a single semester. Long-term peer assessment is based on the Falchikov method, which assesses the project and the process separately. Long-term peer assessment is used to spread a student's peer assessment across several projects, where their work is assessed by several different groups of students. Long-term peer assessment seeks to lower many of the risks identified above, and in doing so provides a more accurate measure of the quality of a student's group work throughout a semester.

Methodology

Selection of Peer Assessment Criteria

"Appropriate evaluation criteria must be established before using peer evaluation as a method to better determine individual efforts within a group" (Beatty, Haas, and Sciglimpaglia 1996).

The peer assessment criteria that were used in this study were devised in a workshop of University of Canberra Television Production students in 1997. The process for selecting the peer assessment criteria utilised a consensual group decision making process to derive a holistic assessment instrument for assessing the participation and contribution of group members, as suggested by Lejk (1996).

In an open classroom discussion, where all the members of the cohort were present,

students brainstormed, debated and secretly voted on the various criteria that made up the peer assessment instrument.

The students eventually settled on 13 criteria, which were expressed in their own words. The criteria were then written up into a document that became the peer assessment instrument. The instrument was given back to the students for crosschecking and approval, which they did without any modifications. The instrument they created was used every semester between the years 1997 and 1999.

In 2000 the student cohort reviewed the peer assessment instrument, and although essentially endorsing it, after much discussion and debate they instigated the modifications detailed in Table 1.

Table 1:
Changes to Criteria in 2000

Criteria Number	Old Criteria	Revised Criteria
6/13	Suggested ideas without arguing	Suggested ideas without negative argument
9/13	Was always enthusiastic and eager to participate	Always turned up on time
10/13	Was honest and dependable	Was honest, reliable and dependable
12/13	Was punctual to all group meetings and shoots	Was punctual to group meetings and shoots, and always responded to phone messages

What is interesting to note about the changes made by the 2000 cohort is that all of them relate to how students function as group members, rather than the quality of a student's contribution to the project. The first change identified that there is a difference between the passionate debating of ideas and being argumentative. The students in 2000 felt that the existing criteria, 'suggested ideas without arguing', might stifle passionate debate. They felt that filmmaking required passion. Changing the phrase to, 'negative argument' encouraged passionate debate, while still frowning upon being argumentative for its own sake, which they felt was disruptive. The other change was to the range of marks, which were expanded as outlined in Table 2.

Table 2:
Changes to Range of Marks in 2000

Old Range of Marks	Revised Range of Marks
16 points - Highest	22 points - Highest
12 points	18 points
8 points	14 points
4 points	10 points
0 points	0 points
-4 points - Lowest	-4 points
	-10 points - Lowest

The students felt that the range of marks was too narrow to cover all the different types of students they were likely to encounter within teams. The students wanted a broader system of marks to reward students whose group work was outstanding, and conversely, punish those students who were lazy, unreliable or disruptive.

These changes support the study conducted by Beatty, Haas, Sciglimpaglia, (1996) who found that, "students expect certain types of behaviour from their group peers". For film and television production students actually turning up to the shoot is one of the most vital expectations they have of each other as group members. For film and television production students, being at the shoot is pretty much everything. A great deal of organisation goes into film shoots, from arranging locations, and permissions, finding

actors, coordinating everyone's availability, organising food and drink, and obtaining equipment on specific days. If after all this careful juggling of numerous stressful variables some students do not turn up, it cannot but have a significant impact on the amount of work, and the quality of work, that is completed.

The Peer Assessment Instrument

"The premise is that group members are in the best position to judge the performances of their peers", Beatty, Haas, Sciglimpaglia, (1996). The instrument presented in Appendix A illustrates the current version of the peer assessment instrument. There are several differences between this instrument and the one designed by Beatty, Haas, Sciglimpaglia, [BHS] (1996). The BHS instrument used one page for each group member. This instrument enables a student to list all the group members, and their self-assessment on the one page. This approach not only reduces the workload for the student, they only have to fill in one form rather than one for each member of the group and themselves, it also reduces the workload for the teacher, and the amount of paper used. Furthermore it enables the student to easily look at the entire group as a whole, making it easier for them to reflect on the range of marks they give. The BHS instrument uses nine criteria; this instrument uses thirteen, and although there are differences in wording, and the number of criteria, the two sets of criteria are similar in essence.

This study corroborates the findings of BHS's 1996 study, which found that such a peer assessment instrument was simple to administer, easily understood by students, including international students, and well accepted by the students. The students also understood that the purpose of the instrument was to assess themselves against their peers, reward those students who worked hard and contributed well, and penalize lazy, uncooperative or disruptive students.

Administering the Instrument

At the beginning of the semester students were each given a copy of the instrument, told how it was devised, and how they should use it. They were told that its purpose was to help allocate a mark for group participation and contribution, because assessing the process is as important as assessing the project. The students were informed that it was a completely separate piece of assessment to the assessment of the project. They were told that each group member would receive the same mark for the project, because there was no way the teacher could genuinely gauge individual contribution. Hence, peer assessment was the method by which they would receive an individual mark for their work on the project. They were told that the peer assessment would remain totally confidential, as suggested by Goldfinch and Raeside (1990), and used solely for deriving an individual mark for participation and contribution. Furthermore, it was made clear to them that they would work on three projects; each time with a different group of students; that they would be peer-assessed by each member of all three groups; and they would also self-assess for each project. This process has been applied from 1998 to 2006.

Formation of Groups

Prior to 1998 students self selected their groups. This resulted in groups being formed based on pre-existing friendships, or estimations of some students' capacity as filmmakers. "Some students chose team members based on their perceptions of their fellow students' expertise and motivation" (Colbeck, Campbell and Bjorklund 2000). Bacon, Stewart and Stewart-Belle's 1998 research proved consistent with Dommeyer's 1986 research that showed that students of similar ability tended to want to work together. In my experience groups formed by friends were not necessarily more harmonious. In fact, some of the most dysfunctional groups began as groups of friends, and ended with very deep rifts in the friendship.

However, the most concerning aspects of allowing students to self selected their groups were that those students who were perceived to be not such good filmmakers, or were

new to the course – such as international students - or were shy students, were often left groupless. Even worse, everyone in the class quickly became aware that these students had not been invited to join a group. This was embarrassing for those students not selected, and awkward for those students who did not want anyone else joining their group.

Consequently, in 1998 I trialled assigning students to groups as suggested by the research of Fiechtner & Davis (1984-1985); Mills & Cattel, Jr. (1998). In their research Colbeck, Campbell and Bjorklund (2000) also supported this approach; however this approach goes against the finds of Bacon, Stewart and Stewart-Belle (1998), which suggests that self-selection has some advantages for students. At the beginning of the semester students were informed that they were being assigned to three different groups, each with a different set of students, and that they would have no say into which group they would be assigned, nor with which students they would work.

Students were formed into six groups for each project: two groups had six students, and four groups had five. These group sizes are similar to those used by Beatty, Haas, Sciglimpaglia, (1996) in their study, and appear to be effective sizes for small student film and television productions. Apart from maintaining as equal a gender balance as possible across the six groups, the allocation of students was a random process. Whether achieving a gender balance is necessary was not tested, although in their study of marketing students Bacon, Stewart and Stewart-Belle (1998) found that, "Apparently, teams of all women, all men, or mixtures of men and women perform about the same". However, Rosser (1998), and Colbeck, Campbell and Bjorklund (2000) suggest that all student teams that include minority group members should include more than one minority group member.

For the second and third project every effort was made to ensure that no student worked with the same student again. Because of the cohort size in some years this has not been possible, however, no student has ever worked with the same student more than twice in a semester: this only ever effected about one or two students. Of course, by second semester repeat grouping is unavoidable. On average, each student worked with 12 different students during the semester. Hence, each student was assessed, for participation and contribution, an average of 15 times—12 peer assessments and 3 self assessments.

The Assessment Components In The Case Study - 1998

There were six assessable components in the subject. Three of these were group-based projects (short films). The details and weighting's of the assessment components are presented in Table 3.

Table 3:

Assessment Components in 1998

1)	Project 1 - Character Portrayal	Worth: 15%
2)	Project 2 - TV Commercial	Worth: 15%
3)	Project 3 - Either Narrative Project or Documentary Project	Worth: 25%
4)	Test	Worth: 15%
5)	Group Participation & Contribution	Worth: 30% Peer Assessed
6)	Project Report This was a written report, which accompanied each of the above projects. The assessment was pass or fail. If the report received a <i>pass</i> the student's grade for the project remained unchanged. If the report received a <i>fail</i> the student's grade for the project was down graded by two increments, e.g. Cr become a P+.	

CIRCUMSTANCES OF THE 1998 STUDY

The project marks from each of the three projects were distributed equally among all group members. At no time, however, even after the final grades were posted, were the

peer assessment marks divulged. If the peer assessment mark had been included as a factor in each student's project mark, as occurs in the Goldfinch method, it would have been easy for students to deduce whether their peers had assessed them positively or negatively. This could have undermined the cohesiveness of the cohort. At the end of each project the peer assessment instrument was handed out. Students completed the peer assessment instrument under exam conditions during lecture time, as suggested by Goldfinch & Raeside (1990). In practice however, this was problematic. Not all students attend lectures, and even fewer students attend lectures that are used for other purposes. Hence, several students had to be chased to complete their peer assessment instrument.

This occurred for many years, until finally in 2007 students were informed that not completing the peer assessment instrument negatively affected the other members in their group, because those students did not receive the full range of peer assessment marks. Consequently, if a student did not complete, and submit, a peer assessment instrument they would lose 20% of their participation and contribution mark for every peer assessment instrument they did not submit. Since introducing this requirement no student has failed to submit a peer assessment instrument.

Calculating the Peer Assessment Mark

In an attempt to neutralise any overly-negative, or overly-generous marks, and achieve a less distorted peer assessment mark, the highest and lowest peer assessment marks for each student were removed.

Ultimately, adjusting the marks in this way proved to be an unnecessary exercise. With the marks unadjusted, the class average was 12.12 out of a possible 16. With the marks adjusted the class average rose to 12.25. For the student who received the lowest mark 7.56, their mark was increased by +0.23, while for the student who received the highest mark 15.38, their mark was increased by +0.34. In all, only 4 students marks went down, these were by, -1.14, -0.06, -0.04, and -0.03. The marks for 19 other students were increased by less than +0.22. The highest increases were by, + 0.34, +0.46 and +0.57.

Since applying this tabulation technique did little more than increase the assessment workload a simple average was used to determine the peer assessment mark.

A student's average peer assessment mark was then converted into a percentage out of 16 (16 being the highest possible mark on the instrument), and that percentage was used to calculate the peer assessment mark out of 30%. Thus, student Z's final peer assessment mark was $(\text{Average}/16) \times 30$.

Results

Assigning Students to Groups

Group discord was still evident, but far less than in previous years. Surprisingly, many students approved of being assigned to groups. By the end of the semester they recognised the value of working with students they had not previously known. They had clearly benefited from forming new working relationships and friendships, and this in turn helped foster a more cohesive cohort of students than had been the case in previous years.

This supports research conducted by Fiechtner and Davis (1984-85); Mills and Cottell, Jr, 1998 and Colbeck, Campbell and Bjorklund (2000) who found that when teachers assign students to groups there is an improvement in performance and satisfaction for the majority of students. At the end of 1998, after being assigned to five different groups (three in first semester and two in second semester), the students were surveyed. Appendix B contains the full survey. The following two questions are extracts from the full survey:

'I would have preferred to have some say in who was in my groups'

Agreed	Neutral	Disagreed
41%	14%	45%

'I would have preferred to stay in the one group all semester'

Agreed	Neutral	Disagreed
17%	14%	69%

Although the number of students who preferred being assigned only surpassed those students who wished to have some say as to who was in their group by 4%, the interesting result is the 14% who were indifferent to self-selecting. The result to the second question clearly shows that students saw advantages to being reformed into different groups for each project. One could surmise that the 14% of students who felt neutral about both questions might have been the same students.

Based on these results students have been placed into groups ever since. Periodically I have surveyed students again, always asking the same two questions, and each year the results have been similar to those obtained in 1998. Overall the majority of students prefer to be assigned to groups. Some students still request to work with friends, and obviously not everyone is happy about being placed in a group with students known to be lazy or uncommitted. However, overall the random assigning of students to groups has proven to be the most democratic and successful method of organising students into production groups. One of the benefits of this method is that generally the commitment of lazy or indifferent students usually improves as the semester progresses.

At the beginning of second semester students are given the option of nominating three people they would most like to work with and three people they would prefer not to work with. However, it is made very clear that in no way can their request be guaranteed. Usually only about one-third of students take up this option to nominate. Of this group of students the majority tend to be the better students who have been developing a project and wish to work together. Occasionally, some students have nominated one or two students they do not wish to work with, and these are usually the students who achieved poorly in the first semester peer assessments. Interestingly, the vast majority of students do not avail themselves of this option, and are happy to work with whoever they are grouped. This may be evidence of the cohesive bonding that random assigning achieves in first semester.

Peer Assessment Marks

Self-assessment is part of the peer assessment instrument, and it was not uncommon for students to self-assess more harshly than their peers assessed them. This was unexpected, and goes against the findings of Boud and Tyree (1995) who found that 'students tend to rate themselves more favourably than they are rated by their peers'. Typically a student might give themselves 18/22 or even 14/22 while awarding other students in their group 20/22 or 22/22. This demonstrates that some students did take the peer assessment seriously. Table 4 summarises the long-term peer assessment marks for three students.

Table 4:

Summary of peer assessments for selected students in 1998

Student	1 st Project	2 nd Project	3 rd Project	Total	Average	Mark x/30%
A: Highest score	12 16 16 16 16 16	16 16 16 16 16	10 16 16 16 16	246	15.38	28.83
B: Average score	4 16 12 12 12	12 16 16 12	8 14 12 12 12	182	12.14	22.71%
C: Lowest score	5 16 12 8 8 8	0 12 8 8 8 8	-4 0 12 8 4 0	121	7.56	14.17

By using peer assessment as part of the assessment process 1 student Failed, 16 received Passes, 12 received Credits, and 3 received Distinctions. The spread of final grades were: 1 mark at 47%, 16 marks between 51 to 64%, 12 marks between 65 to 70%, 2 marks of 75%, and 1 mark of 78%. The single student who did fail failed because they did very badly in the test.

To check what would have happened if peer assessment was not used the 30% of semester marks assigned to peer assessment were re-distributed equally across the three projects. This increased the value of each project by 10%. Each student's mark was then recalculated. The results were surprising. Without peer assessment, 4 students Failed, 23 received Pass grades, and 5 received Credit grades. There were no grades above Credit. Excluding peer assessment also impacted on the spread of final grades: there were 4 marks between 42 to 49%, 23 marks between 50 to 63%, and 5 marks between 65 to 67%.

The case of student E

Student E is an example of how the peer assessment affected an individual case. Student E received the following peer assessment marks over the semester (Table 5).

Table 5:

Summary of peer assessment for student E First Semester 1998

	1 st Project	2 nd Project	3 rd Project	Total	Average	Mark x/30%
Student E	10 16 16 12 16 16	8 16 16 16 9	8 16 16 16 12 9	228	13.41	25.14

Table 6 below details all of Student E's marks across all assessment areas, including the Peer Assessment mark of 30%.

Table 6:

Peer Assessment of Marks for Student E First Semester 1998

1 st Project (15%)	2 nd Project (15%)	3 rd Project (25%)	Test mark (15%)	Peer mark (30%)	Final grade (100%)
7	8	10	11	25	61

If no peer assessment mark had been used, and the 30% allocated to peer assessment was re-distributed equally over the three projects, the final grade would have been 52 as presented in Table 7.

Table 7:

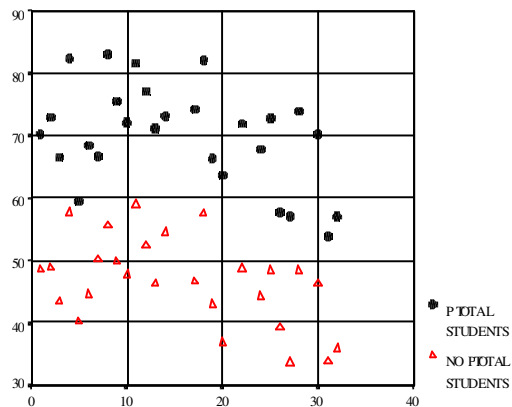
Assessment Marks for Teams containing Student E First Semester 1998

1 st Project (25%)	2 nd Project (25%)	3 rd Project (35%)	Test mark (15%)	Peer mark (0%)	Final grade (100%)
12	13	16	11	0	52

Student E's mark went from 61 with peer assessment to 52 without peer assessment. A further example is student D, who received the highest overall peer assessment mark, 29/30, but was in two poor performing groups for the first two projects, for which student D received 7/15, and 8/15. Without the peer assessment D's final mark would have been 67, equivalent to a Credit. With the peer assessment D's final mark was 78, equivalent to a Distinction.

The scatter plot (Figure 1) below shows each student's final grades for the semester. The P TOTAL includes the peer assessment. The NO P TOTAL is the result of redistributing the 30% peer assessment across the three projects.

Figure 1:
Peer Assessment vs Non-Peer Assessment Student D 1998



The effect of including the peer assessment provides a distinct difference in student grades, and is significant in rewarding individual effort. Long-term peer assessment provided a broader range of marks across the student body, both in terms of the breadth of grades and the scale of grades.

These results have been replicated in subsequent years. See Tables 8, 9 and 10 below. The tables show the range of peer assessment marks awarded to individual students, over subsequent years. These tables confirm that it is not uncommon for a student to receive a harsh peer assessment in one group, but not in subsequent groups.

Table 8:
2000

	Group 1				Group 2				Group 3		
A	16	22	16	16	13	18	16	18	22	10	10
B	-4	10	16	-4	14	16	18	20	22	10	16
C	22	22	22	22	22	18	20	22	22	22	22

Table 9:
2004

	Group 1				Group 2				Group 3				
A	-4	14	14	14	10	14	19	20	18	22	18	20	18
B	21	14	21	18	20	22	14	22	14	14	18	18	18
C	21	22	14	22	22	22	22	22	22	18	22	22	21

Table 10:
2006

	Group 1				Group 2				Group 3				
A	-4	-4	10	18	22	14	21	10	11	10	22	18	22
B	5	22	14	18	14	18	14	14	9	16	22	12	22
C	22	22	22	22	22	20	22	22	22	22	22	22	22

Review of the Results

Year after year, from 1998 through to 2006, long-term peer assessment proved to be a useful instrument for deriving an individual mark for group-work participation and contribution, and providing a workable spread of results. It rewarded students with commitment and passion, and consequently, addressed students' primary complaint with group work—their desire to receive an individual mark that more accurately reflected the work they contributed to the project.

Long-term peer assessment also went some way towards correcting the negative marks that some students received because of personality or goal differences. When comparing

most students' results across all three projects it becomes apparent that there were groups where certain students were acknowledged as helpful members while others were they were perceived in a negative light. The methodology employed in this study makes it difficult to ascertain precisely the reasons why a student received a poor peer assessment mark in one group, and a good peer assessment mark in another group. It is tempting to surmise that this is due to personality and goal differences, but other than teacher instinct there is little hard data to back this up. Because a student's life is inherently complex, there could be any number of external and internal factors that impact on their opportunity to commit to a group film production project to everyone's satisfaction. Similarly, student X might mark student Y harshly, not because student Y did not sufficiently contribute to the project, but because student X had such high expectations that student Y could not possibly live up to them.

The peer assessment instrument used in this study could not provide such data. However, the long-term peer assessment method does go somewhat towards smoothing out such mitigating circumstances.

It is interesting to note that having been told that the peer assessment was completely confidential only very rarely have students requested to see their peer assessment marks. Usually, these are students whose final grades were poor, and they were seeking a justification, apart from themselves, for why they performed poorly or failed. Even in these circumstances the marks were never revealed in detail. Students were only ever given a general statement about their peer assessment grade.

Furthermore, it appears that having been told that long-term peer assessment would be used as an instrument for allocating an individual mark that is comparable to the student's level of commitment, participation and contribution to a project, students appear to readily accept this assurance, and it goes a long way towards satisfying their desire for more individual assessment. Also, students seem content that the peer assessment instrument provides them with sufficient means of ensuring that those students who participate or contribute less are in some way penalized. Even though students never discover by how much a lazy or disruptive student's mark is effected, simply having submitted the peer assessment instrument appears to satisfy their desire for ensuring those students are penalized in tangible way.

Weighting of Peer Assessment Mark

In the 1998 study the peer assessment mark was worth 30% of the final semester mark. Initially this was seen as an appropriate weight. However, it soon became evident that this was too high a proportion of the final semester mark. The typical student is keen and passionate about filmmaking. And although long-term peer assessment was providing a workable spread of results, allocating almost one-third of the semester mark to peer assessment meant that there was an unacceptable inflation of the final semester mark. By 2000 the distribution of marks across the six assessable components had been adjusted. Firstly the Project Report was abandoned at the end of 1999. It was an unnecessary and essentially ineffective assessment instrument. It significantly increased the teaching and student workload while providing very little impact to student grades. These days such a mechanism could be provided by a web blog, but there seems little need for such a device to be re-introduced. Table 11 below outlines additional changes to the overall assessment weightings.

Table 11:
Changes of Assessment Component Weightings

		1998-1999	2000-2006
1)	Project 1 - Character Portrayal	Worth: 15%	Worth: 15%
2)	Project 2 - TV Commercial	Worth: 15%	Worth: 10%
3)	Project 3 - Either Narrative Project or Documentary Project	Worth: 25%	Worth: 30%
4)	Test	Worth: 15%	Worth: 25%
5)	Group Participation & Contribution	Worth: 30% Peer Assessed	Worth: 20% Peer Assessed

Marks were taken away from the peer assessment and the second project and added to the test and the third project. This was done for various reasons. Firstly, removing 10% from the peer assessment better balanced the amount of marks a student received from that component against the other assessment components. Secondly, increasing the value of the third project enabled the students to feel that they were working towards a significant final project, which improved moral. Thirdly, increasing the value of the test still further expanded the range of final semester grades. The spread of grades from a test are likely to be broader than most other assessment tools.

The adjustments implemented are in line with the recommendations of Beatty, Haas, Sciglimpaglia (1996) who suggest that, "peer evaluations, therefore, should constitute 25% of a student's final grade in the class" (p27).

Conclusion

The results of this study are in line with the research of Beatty, Haas, Sciglimpaglia (1996) who found that, "the problem of inequitable contributions can be lessened with a grading system that gives appropriate weight to both individual contributions and group achievement" (p27).

Long-term peer assessment has over the years proven to be a useful instrument for assessing the group-work aspect of student film and television productions. Long-term peer assessment has delivered marks that, when combined with other forms of assessment, such as three film projects and a test, provided a workable spread of grades across the student cohort. The instrument is simple to deliver and manage and students can easily understand its purpose and the process. It performs well at smoothing out harsh peer assessments that can result from personality or goal clashes. For students this method of peer assessment seems to provide them with a satisfactory mechanism to either reward group members who perform well, and, potentially more importantly, to appropriately penalize group members who under perform.

The instrument was designed to be used explicitly for the long-term peer assessment evaluation of an individual's participation and contribution, and does not enable students to assess the quality of their fellow team members' production work. The focus is only on assessing them as group members. More research needs to be undertaken to devise a simple instrument so that students are provided with the opportunity to also assess the quality of their group member's production work. This would not only provide additional assessment information for teachers, but it might also motivate students to consider that the quality of what they contribute is as important as simply turning up to the shoot.

Finally, long-term peer assessment is a useful means of instilling in students the understanding that process is as important as product, and that significant learning does occur during the production process. Being able to perform in accordance with professional practices, and procedures, are highly desirable qualities in university graduates.

References

- Bacon Donald R, Stewart Kim A and Stewart-Belle Sue (1998), 'Exploring Predictors of Student Team Project Performance', *Journal of Marketing Education*, v20, n1, pp. 63-71
- Batra Madan M, Walvoord Barbara E and Krishnan Krish S (1997), 'Effective Pedagogy for Student-Team Projects', *Journal of Marketing Education*, v19, n2, pp. 26-42
- Beatty James R, Haas Robert W and Sciglimpaglia Donald (1996), 'Using Peer Evaluations to Assess Individual Performances in Group Class Projects', *Journal of Marketing Education*, v18, Summer, pp. 17-27
- Boud, D., and Tyree, A. 1995, *Enhancing Learning Through Self Assessment*, Kogan Page, London.
- Brindley, C, and Scofield, S. (1998), The third Year Syndrome: the familiarity that breeds contempt, *Occasional Paper Series*, Number 3, Manchester Metropolitan University
- Brown, S and Knight, P. (1994), The Third Year Syndrome: the familiarity that breeds contempt, *Occasional Paper Series*, Number 3, Manchester Metropolitan University
- Brown, S. (1996), Assessing Individual Performance on Group Projects, *Focus on Learning Problems in Mathematics*, Fall Edition, Vol. 18, No. 4
- Chapman Kenneth J and Van Auken Stuart (2001), 'Creating Positive Group Project Experiences: An Examination of the Role of the Instructor on Students' Perceptions of Group Projects', *Journal of Marketing Education*, v23, n2, August, pp. 117-127
- Colbeck Carol L, Campbell Susan E and Bjorklund Stefani A (2000), 'Grouping in the Dark', *The Journal of Higher Education*, Columbus, Jan/Feb, v71, issue 1, pp. 60-83.
- Conway, R., Kember, D. Sivan, A. and Wu, M. (1993), Peer assessment of an individual's contribution to a group project, *Assessment and Evaluation in Higher Education*, Vol. 18(1).
- Dommeyer, Curt J. (1986). A Comparison of the Individual Proposal and the Team Project in the Marketing Research Course. *Journal of Marketing Education*, 8 (Spring): 30-38.
- Earl, S.E. (1986), Staff and peer assessment - measuring an individual's contribution to group performance, *Assessment and Evaluation in Higher Education*, Vol. 11(1).
- Falchikov, N. (1991), Group Process analysis: self and peer assessment of working together in a group, in S. Brown and P. Dove, P. (eds) *Self and peer assessment*, Paper No. 63, Standing Conference on Educational development, Birmingham, SCED Publications.
- Fiechtner, S.B., & Davis, E. A. (1984-1985). Why some groups fail: A survey of students' experiences with learning groups. *Organisational Behaviour Teaching Review*, 9, 58-71.
- Forman, Janis, and Patricia Katsky (1986). The Group Report: A Problem in Small Group or Writing Processes?. *Journal of Business Communication*, 23 (Fall): 22-25.
- Goldfinch, J, and Raeside, R. (1990), Developing of a Peer Assessment Technique for Obtaining Individual Marks on A Group Project, *Assessment and Evaluation in Higher Education*, Vol. 15(3).
- Goldfinch, J. (1994), Further Developments in Peer Assessment of Group Projects, *Assessment and Evaluation in Higher Education*, Vol. 19(1).
- Grant, A. (1994), Group-project Work: Two Enterprise Case Studies, in L. Thorley, and R. Gregory (eds) *Using Group-based Learning in Higher Education*, Kogan Page, London
- Habershaw, S., Gibbs, G. and Habershaw, T. (1993), *53 Interesting Ways to Assess Your Students*, Melksham, The Cromwell Press
- Hindle, B.P. (1993), The 'Project': putting student-controlled, small-group work and transferable skills at the core of a geography course, *Journal of Geography in Higher Education*, 17(1), pp. 11-20
- Lejk, M, Wyvill, M., and Farrow, S. (1996), A Survey of Methods of Deriving Individual Grades from Group Assessments, *Assessment and Evaluation in Higher Education*, Vol. 21(3)
- Lejk, M, Wyvill, M. and Farrow, S. (1996), A Survey of Methods of Deriving Individual Grades from Group Assessments, *Assessment and Evaluation in Higher Education*, Vol. 21(3).
- Lembke, Svan, and Marie G. Wilson. (1998), Putting the "team" into team-work: Alternative theoretical contributions for contemporary management practice. *Human Relations* 51 (7): 927-945.
- Lerner, Linda D. (1995), Marketing student groups work. *Journal of Management Education* 19 (1): 123-125.
- Mello, J. A. (1993), Improving individual member accountability in small group settings, *Journal of Management Education*, Vol. 17(2).
- Williams, David L., John D. Beard, and Jone Rymer (1991). Team Projects: Schieving Their Full Potential. *Journal of Marketing Education*, 13 (Summer): 45-53.

Appendix A

Exhibit 1

Peer and Self-Assessment Instrument 2000-2007

Please use the criteria below to assess each member of your group, including yourself, for the project you have just completed.

1. Accepted group decisions without ill feelings
2. Had enthusiasm for the project and for the work they were assigned to do
3. Helped the group to function as a team
4. Performed the work they were assigned to do efficiently and competently
5. Stayed within their crew role and did not interfere with other group members' crew role
6. Suggested ideas without negative argument
7. Understood what was required and got on with the job
8. Was actively involved in decision making
9. Always turned up on time
10. Was honest, reliable and dependable
11. Was involved in organising the group and ensuring things got done
12. Was punctual to all group meetings and shoots, and always responded to phone messages
13. Worked harmoniously with other group members and did their fair share of the work

Please use the following scale as a guide to assess each member of your group, including yourself:

- | | |
|---|------------|
| ▪ Successfully achieved in 12 to 13 criteria areas | 22 points |
| ▪ Successfully achieved in 10 to 11 criteria areas | 18 points |
| ▪ Successfully achieved in 8 to 9 criteria areas | 14 points |
| ▪ Successfully achieved in 6 to 7 criteria areas | 10 points |
| ▪ Successfully achieved in 4 to 5 criteria areas | 0 points |
| ▪ Successfully achieved in 3 or less criteria areas | -4 points |
| ▪ Was a hindrance to the group | -10 points |

Name _____	Mark _____
Name _____	Mark _____
Name _____	Mark _____
Name _____	Mark _____
Name _____	Mark _____
Yourself (name) _____	Mark _____

Appendix B

At the end of the 1998 a survey was conducted to obtain student feedback on the study. The survey was designed in three stages. A series of questions were designed to elicit student feedback on the peer assessment process. The questions were given to the Centre for the Enhancement of Learning, Teaching and Scholarship (CELTS), at the University of Canberra, from whom I received feedback on their design and structure. The questions were then modified based on that feedback.

The survey consisted of a series of statements to which students responded on a three-point scale of—agree, disagree or neutral. The results are summarised below. The survey was given to students at the beginning of the last lecture of the semester, and was completed under exam conditions.

1998 End of Year Survey

- 'Peer assessment is necessary for assessing student group work'

Agreed	Neutral	Disagreed
72.41%	20.69%	6.90%
- 'It is useful for students to learn to assess their peers'

Agreed	Neutral	Disagreed
75.86%	10.34%	13.79%
- 'The peer evaluation questionnaire is an effective way of assessing groups work'

Agreed	Neutral	Disagreed
55.17%	10.34%	34.48%
- 'Group marks should be allocated according to individual merit'

Agreed	Neutral	Disagreed
72.41%	20.69%	3.45%
- 'It is important to me that I receive an individual mark, for my group project, that reflects the work I put into it'

Agreed	Neutral	Disagreed
89.66%	6.90%	3.45%
- 'I think everyone in the group should receive the same mark for a group project'

Agreed	Neutral	Disagreed
37.93%	10.34%	51.72%
- 'I expect that the members of my groups will assess me on the work I did and not bring external factors into consideration'

Agreed	Neutral	Disagreed
51.72%	6.90%	41.38%
- 'I expect students in my group will allow friendships and other personal matters to influence their marking'

Agreed	Neutral	Disagreed
65.52%	6.90%	27.59%
- 'Fear of receiving a negative peer assessment stopped me from arguing for my own ideas'

Agreed	Neutral	Disagreed
62.07%	17.24%	20.69%
- 'Peer assessment encouraged my groups to work more co-operatively'

Agreed	Neutral	Disagreed
37.93%	20.69%	41.38%
- 'Peer assessment encouraged me to become more responsible, and take my work more seriously'

Agreed	Neutral	Disagreed
48.28%	24.14%	27.59%
- 'Peer assessment is training me in skills that will be relevant in my work situation'

Agreed	Neutral	Disagreed
58.62%	20.69%	20.69%