What can go wrong with group work and peer assessment? A case study

Martin J. Pitt*

Chemical and Biological Engineering, University of Sheffield, UK

Abstract: It is generally agreed that group work is essential for the practising engineer and should therefore be part of the degree. Assessment is a necessary evil for both staff and students, but becomes much more challenging for group projects. One solution which has been promoted is for some form of peer assessment in which students rate each other’s contribution.

It is rare for academics to experience group work and assessment from the point of view of the student. However, in an exercise, groups of lecturers were given a group task, and also asked to peer assess. The spectacular failure of one group illustrates what can go wrong in such circumstances, and should be cautionary for those using group and peer assessment.

Keywords: Group work, project, team, peer assessment.

*Correspondence to: Martin J. Pitt, Chemical & Biological Engineering, University of Sheffield, Sheffield S1 3JD, UK, email: m.j.pitt@sheffield.ac.uk

1. INTRODUCTION

This paper reports on what happened when a group of academics were set the kind of task commonly given to students, and managed it rather badly. Looking at the task from the point of view of participants in a failing group may elucidate some sympathy, and some caution in interpreting the contributions of individuals.

1.1 Group working
The Engineering Council requires “Evidence of group working” and “working with others” in two of its four general learning outcomes for accreditation of higher education programmes (Engineering Council, 2004) and academics commonly speak of team work as an essential piece of formative education for professional practice. (This despite the fact that academics are in general lone workers.) It is, however, a truth universally acknowledged that a student expected to work must be in need of assessment. Thus if group work is not assessed some students will not put in much effort or even participate at all.

1.2 Group Assessment
The practical problem is well stated by Wilmot & Crawford (2007), who contend that “Academics who feel comfortable setting examinations and single-person assignments are deterred from setting team assessments because they fear that idle students may benefit from the
efforts of their team-mates or that weaker team members might dilute the efforts of the more diligent.”

Goodhew (2010) gives four methods of assessing group work:
1. The academic assesses the output and gives the same mark to all members.
2. The group members negotiate how the marks are distributed, to give the same overall average as the academic’s assessment.
3. The academic assesses the output and also gives marks according to the performance of the members.
4. The academic assesses the output and the students assess each other’s performance (generally anonymously) causing the marks to be varied by some formula.

Method 4 is a common form of peer assessment. Wilmot and Crawford (2007) state that where students simply assess each other’s performance, this tends to exaggerate differences, but by including the students’ assessment of themselves tends to reduce extremes, and report on software, WebPA to assist in automating the process.

1.3 Other factors
Final assessment is not the only driver for the group, nor what should be considered by academics. Aman et al. (2007) report a study of second and third year teams which were variously: happy + functional; unhappy + functional; unhappy + dysfunctional (there were no happy + dysfunctional), and certain factors (such as whether students knew each other already), plus the varying responses of students to the experience. Lester et al (2006) looked at group dynamics in the light of Myers Briggs psychological types.

2. CASE STUDY

2.1 The situation
At a meeting organised by the Higher Education Academy, the academics were put into groups and given the sort of task often given to students, to produce a poster on the topic of the meeting. This was to be assessed by the meeting organisers, and the marks modified by peer assessment. There were five members in one group, whom I shall call A, B, C, D, and E. This is what happened.

2.2 The Group and the Outsider
When the time came to change from lecture to group activity, student E had to go to the toilet. On return a few minutes later, the group had already started and had adopted roles within that group. Thus the group had formed and E was an outsider.

2.3 Group Roles
When group members have taken on roles and a pecking order, fitting in a new person can be disruptive. Student A had taken the role of leader, with B as essentially a henchman. Perhaps E should have been the leader, but that possibility had passed, unless E was sufficiently aggressive to fight for the role. Student A dominated the conversation and was the sole arbiter of what was written down. Student C made positive attempts to include E, and student D was accepting of E.
However, A tended to brush aside contributions by E. Student C said things like “well, I think we should consider what E is saying.”

The group had been given a briefing sheet. Student A had turned this over and was writing notes on the back. Student E reasonably asked “What do we have to do?” and was given a terse reply by A, with C trying to expand helpfully.

Student E asked “Could I please see the brief?” which was grudgingly given with a sigh. On reading through, it turned out that what the group was doing was not what the brief required, but what A had hastily assumed. Thus it could be claimed that E had done a great service to the group by putting them on the right track. Would this be recognised in the peer assessment?

2.4 Group Activity
There was then some fairly inefficient brainstorming, which largely consisted of A putting down A’s ideas plus some contributions from C and a few from D. B mainly seemed to agree with A. Contrary to what is recommended for brainstorming, A did not collect ideas without comment but actively judged them, and did not make an effort to get everyone to speak without interruption.

Student C was probably the most knowledgeable of the group, but this was not fully utilised. C was, however, sufficiently assertive and diplomatic to manage to participate. Student D had good ideas but was less assertive and thus managed to say very little. Student E was more assertive than D and thus made comments and raised queries, but these were largely not followed, and the queries treated with some hostility by A.

2.5 The Poster
The organisers reminded the groups that the aim of the activity was to create a poster, at which point Student A passed the poster sheet and pens to B and said that B would do the poster. There was no discussion. As it happens, student E was a capable artist, who would have admitted this if asked. Perhaps C and D were as well. The evidence suggests that B was not. The supposed poster was just a list of points dictated by A.

Other groups had arrows, blocks and even stick figures with some attempt at design.

2.6 The Assessment
The judges gave marks as they would for a student effort and (quite rightly) what had been produced by this group got a failing mark. However, this was not the end, because the participants had to peer assess. When these were collected, the result was that student A had contributed most and student E the least. Thus A overall scraped a passing mark while E got the lowest mark in the class.

2.7 Comments
Two points must be made. Firstly, this group consisted of experienced academics, familiar with student group work, not first year students. If they could make such a mess of it, what chance do some students have? Secondly, the marks awarded had no consequence for the participants.
However, if student E had really been a student, this failed group activity could have made the difference between honours categories or proceeding to the next year.

Overall, the group could be characterized as unhappy + dysfunctional. While A was happy during the group activity, A was unhappy when the posters were put up and even more unhappy when a poor mark resulted.

2.8 Another case
In a different lecturer’s workshop (another place, another time) the group also had to make a poster. The following happened. There were 4 academics playing the part of students.

Student F suggested making a mind-map. Student G agreed enthusiastically, Student H agreed and Student J said he did not like mind-maps. Essentially Students F and G made the mind-map and developed it into a poster on the topic provided, but with significant participation by H. A poster which was generally agreed to be very good was produced, but Student J refused to have his name on it.

3. DISCUSSION

3.1 A Confession
Let me first admit my prejudice. I did not believe that peer assessment was a good thing. Nevertheless, I was prepared to be persuaded and went with an open mind to a meeting to promote this practice, providing support and methods. Some students were also produced, who felt that it had been a good thing, and academics eagerly related the sometimes complex methods they used. The point of the exercise described above was to show the virtues of one method.

3.2 What might be wrong with peer assessment?
My objections (before this exercise) were mainly as follows:

1. Assessment is the job of the academic.
2. A student’s progression or degree should not ever depend on other students’ opinion of them.
3. It is very easy for one or two students to be outsiders who may find it difficult to participate and in addition be marked down.
4. There may be individual antipathies, or conversely students wishing to find favour with another.
5. An individual student may have a problem (such as a recent family tragedy, or illness) not known to the others.
6. It gives the possibility of strategic game-playing, particularly by the most capable student to maximize marks at others’ expense (Pitt, 2000).
7. Advocates often explain that it only makes a small contribution (typically 5%) to an overall module mark. If it is so good, why is it limited?

3.3 What might go wrong with a group project?
Both the psychological literature and management books (e.g. Wilke & Meertens, 1994; Robbins & Finley, 1997) describe the complex interplay of group work and how it can go wrong. The classic way of assigning students to groups is to share out the clever ones and the duffers, one
per group, i.e. to sort on academic performance to date. However, there are other ways people can vary. The educational literature in general and the engineering educational literature deal with several methods of categorizing individual styles of behaviour in groups, and many papers on methods intended to make them work. Lester et al. (2006) used the Myers Brigg categories and found that of their set of engineering students, 64% were SJ (Sensing/Judging), 18% NT (Rationals), 9% NF (Idealists, Feelers) and 9% SP (Sensing/Perceiving). We also have percentages of male/female, home/foreign etc students. In practice some groups work well and some do not, with individual behaviour varying from not attending at all to attending but not contributing at all as extreme behaviours, and confusion and misunderstanding as common features. In addition groups often split into smaller, sometimes competing, groups, typically pairs or trios.

3.4 Case Study Outcomes
The exercises described were time limited to students in a room, so the problem of someone not turning up to group meetings (which happens with longer term projects) was avoided. The group size was small and the task very limited, so there was no reason or excuse to split up. Thus some of the ways of failure open to larger groups on longer projects were unavailable. In addition, all participants were native speakers of English and had some measure of self-confidence. All were associated with engineering, so there was not even a possibility of a split – e.g. with arts and humanities people. (The group was mixed male and female, but did not seem to divide in any way on these lines. I would not expect there to be issues with sex or ethnicity from a group of academics, but it does occur with students.)

Nevertheless, several things went wrong, and the group did not produce the best possible work it could have managed from the talents available. The rapid assignment of roles and the phenomenon of the outsider were a surprise to me. It may seem extreme to suppose that a few minutes could make the difference, but this is supported by work on group behaviour, well described by Harris (1998). There are all sorts of ways in which a student can be an outsider to a group that he or she is supposed to be part of. However, being ill for a day or just arriving late for good reason can be enough. A student transferring from another institution or having taken a year out, coming to a set of people who know each other already will find it difficult to become a full part of the group. This is particularly so with group projects given at the start of the academic year, for students who are shy, and for students coming to Britain who are less confident in spoken English.

I did not view the process by which A became leader. It is my judgment that C would have been better in this role, and would have been a more inclusive facilitator. The group would surely have benefited overall by utilising E’s artistic ability.

4. CONCLUSION
The experience provided an insight into the difficulty and possible unfairness of group working and peer assessment. Because of the multi-parameter nature of the differences between people, I do not believe it is practical to allocate group members in any way which makes all groups equal. Academics should be cautious of allocating marks on a basis of blame for failure, whether the blame comes from perception of the academic or the students themselves.
5. REFERENCES

Aman, A., Poole, G., Dunbar, S., Maijer, D., Hall, R., Taghipour, F., & Berube, P., (2007) Student learning teams: viewpoints of team members, teachers and an observer Engineering Education 2 (1) 2-12


