

## **PEER ASSESSMENT AS AN APPROACH TO JUDGE GROUP WORK: DOES IT WORK?**

**Vicente Rodilla**

Universidad Cardenal Herrera CEU, Dep. Fisiología, Farmacología y Toxicología,  
Área de Toxicología  
Avda. Seminario s/n, 46113 Moncada, Valencia

*Rebut: maig de 2006. Acceptat: desembre de 2006*

### **ABSTRACT**

Peer assessment strategies have been around for quite a while and the benefits derived from their implementation have been described and demonstrated. Peer assessment increases student motivation, builds up on self confidence and stimulates critical thinking and is a particularly useful strategy to assess the contribution of each student to group work. It is however a type of assessment which can be difficult to implement if the students do not support it. In this paper we describe and discuss the relative lack of success we have had when we have tried to use self and peer assessment to evaluate group work in Spain, whereas a very similar strategy worked extremely well for us in the United Kingdom. We also discuss the reasons which may explain this apparent contradiction.

**KEY WORDS:** assessment, peer assessment, group-based assessment, evaluation.

---

### **RESUMEN**

Las estrategias de evaluación que realizan los estudiantes sobre sus propios compañeros (*peer assessment*) se conocen desde hace unos años y se han descrito y demostrado los beneficios que de su utilización se derivan. Estos métodos evaluativos incrementan la motivación de los estudiantes, refuerzan su autoestima y estimulan el razonamiento crítico. Es además una estrategia muy útil para determinar la contribución que realiza cada estudiante cuando se trabaja en grupo. Es sin embargo un tipo de evaluación que puede resultar difícil de implementar si no recibe el apoyo de los estudiantes. En este artículo describimos y discutimos el relativo poco éxito que hemos tenido cuando hemos tratado de utilizar autoevaluación y peer assessment para evaluar el trabajo en grupo en España, mientras que una estrategia similar nos funcionaba

extremadamente bien en el Reino Unido. Discutimos también las razones que podrían explicar esta aparente contradicción.

**PALABRAS CLAVE:** evaluación, peer assessment, evaluación de grupos.

---

## INTRODUCTION

In recent years, social and emotional gains attract as much interest as cognitive gains and to this aim, cooperative and group learning, and peer assessment, play a very important role in a number of educational establishments, from kindergarten to secondary school (Topping, 2005). This is also true for Universities where a relative big weight is being placed on learning and developing transferable skills, without compromising the acquisition of knowledge by the students.

It could be argued that because of the inherent gregarious nature of the human being, teaching and learning in groups is probably a more natural approach than the individual approach until now so prevalent in formal education. It has been shown that cooperative group work improves in the students the acquisition of information and its retention as well as social and communication skills and self-confidence (Johnson et al., 1998). Therefore group work can be a very important teaching strategy in any course, in as much as it helps to develop teamwork skills without sacrificing the acquisition of knowledge. An important question that arises when group or collaborative work is implemented is how to assess the work the students have carried out with fairness. The lecturer can evaluate the final outcome of collaborative work, but it would be impossible for the lecturer to measure and assess the contribution of any individual student within the group. Who is then best placed to carry out a fair and objective assessment of the individual performance of each student within the group? It is certainly not the lecturer, particularly given the fact that much of the work has been undertaken away from the classroom and away from his/her supervision. This assessment is better carried out by the members of the group: only the students know, and can evaluate therefore what has been the contribution of each one of the members of the group to the task assigned.

Assessment is an intrinsic part of any teaching method. It has been stated that much of what happens in university studies is driven by assessment, and it has been estimated that teachers can spend between one third to half of their teaching time dealing with matters which are related directly or indirectly to assessment (Schaffer, 1991). However, effective assessment is more often than not difficult to implement adequately to make it a part of the teaching process by

which student learning is enhanced and reinforced. Sally Brown (1999) maintains that the methods used to assess students in Universities and Colleges around the globe are extremely limited. In fact, she has estimated that approximately 80% of total assessments are traditional exams, essays and reports (Brown, 1999). Most of these more traditional forms of assessment are generally teacher-centred assessments and therefore they limit the opportunity for the students to fully understand educational objectives and furthermore, they limit the opportunity to develop critical evaluation skills by the students. Other less conventional assessment approaches would be self-assessment, peer assessment and group based or computer based assessments. Amongst the different approaches and methods used to assess students, one that has been successfully described by some authors and which is becoming widely used in a number of Universities around the world is peer assessment (Brown & Knight, 1994; Topping, 1998; Billington, 1997; McLaughlin and Simpson, 2004; Kaufman et al., 2000; Elliott and Higgins, 2005). Peer assessment provides feedback opportunities for the students as they are assessing other students. Some papers promote peer assessment usage and several authors have tested its validity and reliability in a number of applications such as student presentations, group work, projects, reviews or essays handed out to the class, etc. (Billington, 1997; Topping, 1998; Topping, 2005).

An intra peer assessment in which the students assess the performance of other students with whom they have been working is particularly useful in as much as it allows the lecturers to explore and obtain important information about group dynamics. If properly performed, intra peer assessment allows the lecturer to detect individual student contribution to the overall product. It may be used to detect leaders in each group, and perhaps more importantly it is invaluable to detect those students who are not pulling their weight or are simple hitchhikers (those who take a free ride at the expense of their more industrious colleagues) (Kaufman et al., 2000). Since one would tend to think that the students would be in favour of a fair and equitable system of assessment for group work and sometimes are concerned that the group work does not reflect their own individual effort, peer and in some cases self- assessment can be the answer. In fact, we would argue that for assessing group work, peer assessment is about the best way for the lecturer to get a fair idea of the contribution each student has made to the overall project.

For several years, we successfully used self and peer assessment to judge the contribution of individual students to group work in a Biomedical Sciences course in the UK. In this paper we describe and discuss the difficulties we have experienced in implementing peer assessment to judge group work with Spanish students. It was never our intention to carry out a controlled experiment by comparing two different populations of students to peer assessment. What we

present here is rather the results which have appeared after several years of using peer assessment to evaluate the group work of our students both in the United Kingdom and in Spain.

## METHODOLOGY

As stated above, we developed a peer assessment strategy for group work assessment which was implemented in three successive academic years (from 1997-98 to 1999-00) at the School of Applied Sciences at the Robert Gordon University (Aberdeen, Scotland) and an assessment pro



SCHOOL OF APPLIED SCIENCES  
AS3054 Environmental Toxicology

### PEER AND SELF-ASSESSMENT OF TEAM WORK.

This form should be completed individually. There is no need to consult other group members on the completion of this assessment form.

Use the following score to mark the degree of your contribution and that of your colleagues to your project:

- 0= None (only to be given in exceptional circumstances)
- 1= Very little
- 2= Limited
- 3= Some
- 4= Significant
- 5= Considerable
- 6= Outstanding

Your name:

Group number:

ASSESSMENT CRITERIA	STUDENT NAME	SELF			
Attendance at group meetings (regularity, punctuality & cooperation)					
Contribution of ideas and suggestions for the project					
Efficiency in carrying out the tasks assigned by the group					
Acceptance of fair share of the work					
Contribution to the preparation of the presentation					
Overall contribution to the project					
<b>TOTAL</b>					

Dr V Rodilla

**Figure 1.** Self and peer assessment pro forma which we used at the Robert Gordon University, Aberdeen, to assess student contribution to group work.

suggestions, efficiency in carrying out the tasks assigned by the group, acceptance of their fair share of work, contribution to the finished project (poster, presentation, report, etc) and their overall contribution to the project. In each category, the student marked themselves and their peers on a scale from 0 to 6. (Figure 1). At the Robert Gordon University peer assessment for

forma was designed based on Stefani and Tariq (1996) (Figure 1). This assessment pro forma was translated into Spanish and used for self and peer assessment when group work was involved at the Faculty of Experimental and Health Sciences at the Universidad Cardenal Herrera CEU (Valencia, Spain) (Figure 2). In both universities, the reasons as to why peer assessment was necessary and the importance of peer assessment to the overall assessment of the group work was explained to the students. Using this assessment form the students had to grade themselves and each member of their group on several aspects of the group work, which were attendance at group meetings, their contribution with ideas and

group work amounted to 10% of the mark for the group assignment. The lecturer (or lecturers) grading of the group work contributed a 50-60% and the remaining 30-40% was obtained either by the students assessing the group assignment or in some cases from the assessment grades given to the group by all the other groups. At the Universidad Cardenal Herrera during the years 2000-01 to 2002-03, when a group exercise was carried out and the attendance to laboratory practicals was approximately a 30% of the final mark. Self and peer-assessment varied from 20 to 40% of the total mark for group assignment. The mark given by the lecturer to the group work had a weight of 60 to 80%.

**AUTOEVALUACIÓN Y EVALUACIÓN A SUS COMPAÑEROS DEL TRABAJO EN GRUPO**

Utilice los siguientes criterios para otorgar la calificación oportuna.

0 = No contribuye (sólo se otorgará en circunstancias excepcionales)  
 1 = Contribución muy limitada  
 2 = Contribución limitada  
 3 = Contribuye  
 4 = Contribución significativa  
 5 = Contribución considerable  
 6 = Contribución inmejorable (Ha hecho prácticamente todo el trabajo).

Criterios	Nombres	Escriba en cada casilla el nombre de un compañero			
	Autoevaluación	Mi nombre: Rafael C.	Rafael C.	Juan José S.	Amparo P.
Participación en reuniones grupo (Regularidad, puntualidad, cooperación)	6	6	6	6	
Contribución al grupo con ideas y sugerencias	4	5	4	5	
Aceptación de su parte proporcional del trabajo	5	5	5	5	
Contribución a la preparación de la presentación	5	4	5	4	
Contribución total al proyecto.	5	5	4	4	
	5	5	4.8	4.8	

**Figure 2.** A used example of a pro forma used at the Universidad Cardenal Herrera CEU to evaluate student contribution group work. The surnames of the students have been deleted.

## RESULTS AND DISCUSSION

When self and peer assessment was used with our Biomedical Science students at the UK, the results were highly satisfactory and allowed us to evaluate group work with fairness. While some students expressed their doubts and reservations about assessing their peers the first time this system was used, particularly if they were considering giving a low mark to any colleagues, those doubts were not an issue in subsequent years, although they had to be reassured that the grades given to any student would be treated confidentially.

On one occasion while teaching on a module on Human Anatomy and Physiology the group work did not involve peer assessment and the comments the students made during the evaluation of the course were very clear (Table 1). The students were clearly against a single mark given equally to each individual within the group and wanted to assess their own contribution.

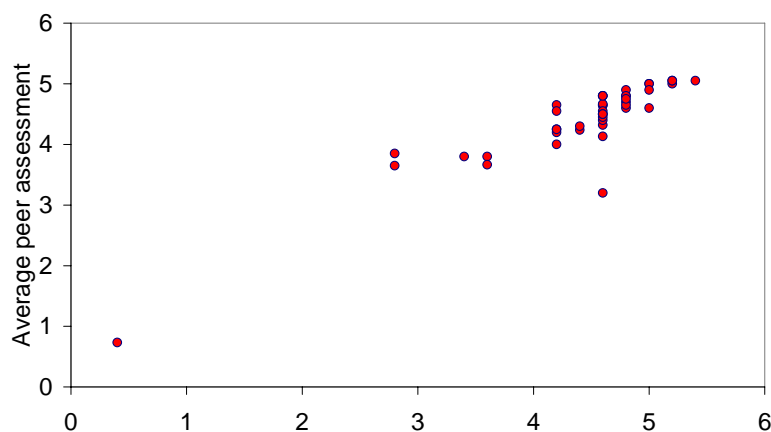
For several years we have tried to put into practice this peer assessment approach with our Pharmacy and Veterinary students at the Universidad Cardenal Herrera but with limited success. The first year we tried to use it, everybody in the class awarded themselves and their peers in the group either a 5 or a 6, which invalidated the assessment, because all students in the group had the same marks. Kaufman and co-workers (2000) have described similar situations, with students agreeing to give one another identical ratings and they see this fact as evidence that the teams are working correctly and everyone participates evenly in the group work. However in their experiment, this happened in 5-10% of the teams, which can be acceptable. We have had a similar situation in Scotland. With such a low incidence, one cannot discard that this is a

**Table 1.** Comments made by students when peer assessment was not used to evaluate their non-standard assessment (NSA).

NSA should be able to be marked, at least in part, on an individual basis
Change the NSA (We must assess our project group)
NSA should be on individual basis, not group.
Students should be marked on their own input/knowledge to the NSA.
Group assessments are very unfair, and not a true reflection on each individual

coincidence because the group is working well and everyone pulls their weight, or whether the students have reached an agreement indeed. In further years we have tried different modifications to make peer assessment work, either by explaining in much more detail the importance of fair assessment and what we wanted to get from the students marks, or by adding further rules to the marking system. For instance, if in the group there are  $n$  students the same mark could only be used  $n-2$  times in each category, but this strategy complicates the assessment unnecessarily.

In both cases and with a number of groups the students had made a prior pact, so that everybody in the group had the same overall mark. When this happens, the whole exercise is pointless, because we (the lecturer) fail to find out, and therefore assess, what the contribution of each student has been to the group project.



**Figure 3.** Scatter plot showing the relation between self and average peer assessment marks for 2002-03 academic year for Toxicology students in the Pharmacy degree.

Nevertheless on some occasion such as the Veterinary (2003-04) or Pharmacy students (2002-03) self and peer assessment was correctly used and with a certain degree of success. For the sake of the discussion we present here the results

for the Pharmacy course taught in 2002-03. In Figure 3 we show the correlation between the self-assessment and the average peer assessment marks on a group exercise. There is a statistically significant positive correlation between the marks the students award themselves and the average awarded to each student by their group colleagues ( $\rho=0.754$ , Spearman rank test;  $p < 0.05$ ). Half the students ( $n= 22$ ) awarded themselves a slightly higher mark than that their colleagues awarded, whereas the opposite, colleagues awarding a higher mark than themselves happened with 17 students (38.6%). With 5 students (11.4%) the grade awarded by colleagues or by themselves was identical. One student got a particularly low mark by their peers, but he was aware of his own contribution and also awarded himself a low mark. These results are similar to those described by other authors (Kaufman et al., 2000). One team was not included in these calculations since it was considered to be a dysfunctional team: three members of the team failed to submit their ratings. Even on these occasions in which self and peer assessment seems to work, the students remain unconvinced about their utility since they did not like the idea of grading themselves and less, grading their peers.

During the year 2004-05, and given the, at least from our point of view, unsatisfactory results obtained in the previous years, a group exercise was put into practice which did not involve peer assessment. Each group of students prepared an 10-15 minute oral presentation on the toxic effects of a selected drug (or a family of compounds) to be given in class to their peers and they also submitted a written report. The lecturer assessed the presentation and the report, grading the group work and allocated each student in the group the same mark irrespective of their contribution. We expected to hear some comments about the fairness/unfairness of this system, and how single student input was not taken into consideration, but this was not the case. In contrast to their British counterparts (see comments on table 1), the Universidad Cardenal

Herrera CEU students seemed to be quite happy sharing the marks irrespectively of their contribution. They certainly did not make a single comment about being unfair or wanting to assess themselves individually.

According to the literature (Kaufman et al, 2000; Felder and Brent, 1994), and in our own experience, the most common complain the lecturer receives when the possibility of group work within the class is mentioned, is about hitchhikers, but when the students know that their mark will effectively lower the mark of that particular student but not lower the marks of all the other working members of the team they cease to complain about the unfairness of group work. However we have not got this type of complain with our Pharmacy or Veterinary students in Spain. In most cases in fact, the students resent having to downgrade a colleague because his/her contribution to group work is not acceptable and they tend to cover for that particular student.

The marks that were obtained from the peer assessment were substantially increased from a 10% of the total mark in Scotland to a 40 % in Spain. This was done because the students in Aberdeen always complained that the proportion of marks allocated to peer assessment was too small. Similar comments have been reported by others who also allocate a 10% to peer marking (Elliott and Higgins, 2005).

The question we ask ourselves is “why did self and peer assessment worked so well in The School of Applied Sciences in Aberdeen and is not well received in Spain?” Felder and Brent (1994) have reported that lecturers who have attempted to install cooperative learning in their courses “frequently encounter resistance and sometimes open hostility from the students”. This is probably something that has happened to us since our Veterinary and Pharmacy students are not as familiar with these types of strategies as their Scottish counterparts. In fact non standard assessments, whether group or individually based, and using a variety of different approaches (posters, seminars, oral presentations, written work, etc) were a common occurrence for most modules (subjects) in a degree or even in the Diploma courses. There are also differences in our students and our Universities, which we have not taken into account. In Scotland, most students either live on-campus or nearby; at Universidad Cardenal Herrera-CEU, many students live off campus, some needing to travel an hour or more to come to college. Spanish timetable is also very different from the timetable in other European countries. Scottish students would be required to be at college during the day from 9:00 h to attend lectures, laboratory practicals, meetings, etc., but their college engagements would be over by 17:30 h at the latest. The number of subjects (or modules) in which the student registers is much more regulated by the



Faculty and the University and as a result, all (or at least most) students enrolled on a course follow either the same or a very similar timetable. The students at Universidad Cardenal Herrera-CEU have a broader timetable with lecturers, practicals and other activities distributed (almost at random) throughout the day. To begin there is the lunch break which can occupy two to two and a half hours in the middle of the day. Additionally, there are practically no restrictions on the subjects the student chooses to study in a particular year. Thus, a student may begin the day at 8:30 with the first lecture and have the last at 19:00 h. In between there are many free hours (one here, two there, another elsewhere) depending on the number of subjects that particular student is registered. Time availability to do group collaborative work in Spain can be therefore much more restricted. All these constrictions can make meetings between students outside lectures, if not impossible at least, very difficult.

There are also educational differences in the students when it comes to assess their peers. Spanish students often question the validity of these exercises, being reluctant to grade themselves and even more reluctant to grade their peers. They tend to see downgrading a student who has not contributed to the group work as a telltale activity; worse still they think the lecturer should not find out. On occasions the group has been discovered to be covering for students who do not pull their weight to the team work. However, the attitude in Scotland also reported by others (Elliot and Higgins, 2005; Felder and Brent, 1994) is very different, with the students having no problems down-grading individuals who were not contributing what they should to the group effort.

Given the demonstrated benefits of cooperative learning coupled with self and peer assessment, by means of which the students are better motivated, tend to show higher academic achievements, achieve better high-level reasoning and critical thinking skills, normally through working in groups with their peers gain a deeper understanding of the learned material, develop their teamwork skills, etc (Felder and Brent, 1994; Elliot and Higgins, 2005, Stefani and Tariq, 1996; Topping 1998, 2005; Johnson et al., 1998; Hanrahan and Isaacs, 2001; Gatfield, 1999) we must continue to try to welcome this type of learning amongst our students.

Perhaps where collaborative learning and peer assessment are difficult to implement, as it is our case, we (the lecturers) should try to make the students our allies by involving them more actively in the design of the assessment forms and the computations for the overall marks for student and group work, so that they can perceive it as a fair system of awarding marks. A number of authors (see Elliot and Higgins, 2005) have demonstrated that self and peer assessment strategies are well received by the students when they are actively involved in them.

## CONCLUSIONS

While it is generally agreed that a diversity of assessment approaches are needed in order to assess students fairly, not all of these strategies are useful in all settings. For instance, peer assessment is not well received by Spanish students when compared to their British counterparts. In our experience, Spanish students are extremely reluctant to participate in self or peer assessment exercises. They tend to see peer assessment as a telltale activity short of being moles or defectors. They also consider it to be an interference in their own world. British students are also far more competitive amongst themselves than Spanish, to the extent that in some cases they do not even share lecture notes. These very different attitudes could be a very important factor which may help to explain why some, successfully proven, teaching strategies do not seem to work with certain groups of students. Perhaps there is a philosophical and theoretical framework which explains and directs the context in which peer assessment takes place and it is this framework and its connotations that needs to be discovered and exploited to make a fuller use of peer assessment.

## REFERENCES

- Billington, H.L. (1997). Poster presentations and peer assessment: novel forms of evaluation and assessment.
- Brown, S. (1999). Promoting Effective learning through assessment. The Robert Gordon University, Aberdeen.
- Brown, S. and Knight, P. (1994). Assessing learners in higher education. Kogan Page, London.
- Gatfield, T. (1999) Examining students satisfaction with group projects and peer assessment. *Assessment and Evaluation in Higher Education*. 24: 375-377.
- Hanrahan S.J. Isaacs, G. 2001. Assessing self and peer assessment: the students' viess. *Higher education Research & Development*. 20: 53-70
- Elliott, N., Higgins, A. (2005). Self and peer assessment – does it make a difference to student group work? *Nurse Education in Practice*, 5: 40-48.
- Felder, R.M., Brent, R. (1994). Cooperative learning in technical courses. ERIC Document Reproduction Service Report ED 377038. (<http://www.ncsu.edu/felder-public/Papers/Coopreport.html>).
- Kaufman, D.B., Felder, R.M. , Fuller, H. (2000). Accounting for individual effort in cooperative learning teams. *J Engr Education* 89: 133-140.
- Johnson, D.W., Johnson, R.T., and Smith, K.A. (1998). Active learning: cooperation in the college classroom. Edina, Interaction book Co. Minnesota.

- McLaughlin, P., Simpson, N. (2004). Peer assessment in first year university: how the students feel. *Studies in Educational Evaluation*, 30: 135-149.
- Schaffer, W.D. (1991). Essential assessment skills in professional education of teachers. *Education Measurement Issues and Practice*, 10: 3-12.
- Stefani, L.AJ. , Tariq, V.N. (1996). Running group projects for firs-year undergraduate students. *J Biol Education* 30: 36-44
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Rev. Educ. Res.* 68: 249-276.
- Topping, K. (2005). Trends in peer learning. *Educational Psychol*, 25: 631-645