

Using peer assessment to evaluate teamwork from a multidisciplinary perspective

Planas-Lladó, A.^a; Feliu, L.^b; Castro, F.^c; Fraguell, R.M.^d; Arbat, G.^e; Pujol, J.^e; Suñol J.J.^f; Daunis-i-Estadella, P.^g

^a *Pedagogy, University of Girona, Girona, Spain*

^b *Chemistry, University of Girona, Girona, Spain*

^c *Computer Science, University of Girona, Girona, Spain*

^d *Geography, University of Girona, Girona, Spain*

^e *Chemical and Agricultural Engineering and Food Technology, University of Girona, Girona, Spain*

^f *Physics, University of Girona, Girona, Spain*

^g *Statistics, University of Girona, Girona, Spain*

* Corresponding author: Planas-Lladó, A, Faculty of Education and Psychology, Department of Pedagogy, University of Girona, Sant Domènec, 17071 Girona, Spain (e-mail: anna.planas@udg.edu)

Short biographical notes of the contributors:

Planas, A. is a Lecturer in the Department of Pedagogy at the University of Girona. She has a background in community development and worked in the voluntary sector and youth and cultural services of local government before joining the University. Her research interests include community development, assessment, and youth and cultural policies. In the educational field, she has published research on quality improvement in education and student participation in university governance.

Feliu, L. obtained a BSc degree in Pharmacy from the University of Barcelona, where she began her scientific career and completed her PhD degree in 2000. In 2001, she obtained a two-year postdoctoral position at the University of Montpellier as Marie Curie Fellow. Following that, she moved to the University of Girona, where she obtained a Marie Curie Reintegration Grant. She joined the Department of Chemistry's LIPPSO group as lecturer and became Associate Professor in 2008. Her main research topics are centered on the design and synthesis of peptides with antimicrobial and antitumor activity.

Castro, F. obtained a PhD in Computer Science (2002) from the Polytechnic University of Catalonia, where he had previously obtained a degree in Computer Science Engineering. He holds the position of Associate Professor at the University of Girona. His current research is focused, on the one hand, on the application of intelligent algorithms to different practical fields, such as medicine and interior design, and, on the other, on simulation of the illumination of interiors involving both natural and artificial light.

Fraguell, R.M. obtained a PhD in Human Geography from the Autonomous University of Barcelona in 1994. She is a Professor of Geography and Tourism Studies at the University of Girona. Her research focuses on spatial analysis in tourism, in particular strategy design for the integrated management and sustainable development of tourist destinations, topics on which she has a number of scientific publications.

Arbat, G. is an Associate Professor at the Department of Chemical and Agricultural Engineering and Food Technology at the University of Girona. His research is focused on the improvement of water and energy efficiencies in irrigation, soil water modeling under different irrigation systems and water re-use in agriculture. He is the author of several scientific papers dealing with these topics.

Pujol, J. is an Associate Professor at the Department of Chemical and Agricultural Engineering and Food Technology at the University of Girona. His research has focused on issues relating to the economy of natural resources, particularly water. He is the author of several scientific papers dealing with these topics.

Suñol, J. J. is Full Professor of Applied Physics at the University of Girona. His field of specialization is the production and thermal and structural characterization of new materials. He has published over 170 scientific articles and more than 200 conference papers. He is coauthor of 8 university physics books and a member of the Physics Education group belonging to the Royal Spanish Society of Physics.

Daunis-i-Estadella, P. is Professor of Statistics at the University of Girona. He obtained a PhD in Statistics and Operational Research from the Polytechnic University of Catalonia in 2005. His research focuses on three areas: the analysis of compositional data, methods of data fusion, and applications of statistics in medical research. He is also a member of the Spanish Statistical Society's education group.

Using peer assessment to evaluate teamwork from a multidisciplinary perspective

Abstract

This article analyses the use of peer evaluation as a tool for evaluating teamwork and students' perceptions of this type of evaluation. A study was conducted of six subjects included on five degree courses at the University of Girona. In all of these subjects, students carried out a team activity, evaluated the performance of the team and the involvement of its different members, and responded to a survey on their perceptions of this evaluation system. We found the main factors influencing the evaluation and perception of teamwork to be teachers' and students' prior experience in this type of evaluation activity, the field where it is applied, the academic year students are enrolled on and the weight of the activity in the final mark. The results show that, in general, students' views regarding such evaluation procedures are positive.

Keywords: peer-assessment, teamwork, higher education,

Peer assessment as a learning strategy

The implementation of the Bologna process has led to new practices of assessment in higher education. Among them, this process stipulates that students must receive both summative and formative types of assessment. Summative assessment (the assessment of learning), in which students have a passive role, focuses on learning as a finished product and is used to grade students. Formative assessment (assessment for learning), on the other hand, is assessment in which students play an active role, facilitating the progressive assimilation and development of content and competences and the detection of acquired learning via evaluation of their own activity and/or that of their peers and the design of strategies and assessment models (Cestone, Levine and Lane 2008).

In order to meet this aim of promoting formative assessment, teaching methodologies involving students in the training process have become more widespread in higher education in general and at the University of Girona in particular. Such methodologies include, among others, self-assessment, peer assessment and coassessment. However, their implementation in the university sphere has been gradual and often experimental in nature (Wen and Tsai 2006; Vickerman 2009; Rodríguez, Ibarra and Garcia 2013), and in Spain they have not been applied in a very thorough manner. Rodríguez, Ibarra and Garcia (2013) have attributed this low implementation to a lack of

favourable institutional policies and argued that only institutions' clear support for these assessment methods can change the approach to the assessment of learning.

The benefits of these modes of assessment have been extensively studied. On the one hand, they favour the learning process and help students to develop a variety of skills (Lindblom-Ylänne, Pihlajamäki, and Kotkas 2006). On the other hand, they enhance not only students' participation and autonomy but also the degree of responsibility they adopt towards their learning process. Other benefits of these types of assessment are improving student motivation, acquiring a greater understanding of content, improving the effectiveness and quality of learning, and increasing the capacity for critical analysis (Topping et al. 2000; Wen and Tsai 2006; Lindblom-Ylänne, Pihlajamäki, and Kotkas 2006; Vickerman 2009; Topping 2009). In line with this, Kearney (2013) describes the results obtained from an experiment based on self-assessment and peer assessment via the *Authentic assessment for sustainable learning (AASL)* model, which is a combination of professor, peer and self-assessment. This author shows how summative assessment tends to easily lead to demotivation and a lack of commitment in students, whereas they tend to become more involved when using formative assessment procedures. This tendency is especially observed when students also have to wholly or partially design the assessment tools.

Peer assessment in teamwork

Learning teamwork skills is key to the development of professional life in many spheres. Hence, teamwork is included on the curricula of many degree courses in the Spanish university system as a cross-disciplinary competence which brings numerous benefits. Among others, it increases the motivation of students, teaches them to take risky decisions, stimulates their creativity, and helps them assimilate new learning (Pallisera et al. 2009).

In order for students to acquire competence in teamwork it is not enough, as often happens, to simply introduce one or more teamwork practice activities on degree courses. To achieve this aim, it is also essential to help students organize, plan, review the work and functioning of the team and propose improvement goals to learn to work as part of a team (Pujolàs 2008). It is therefore necessary to reflect on experiences, the role each student and other team members have adopted, work dynamics, the level of planning, the flow of communication between members, etc. Unfortunately, this subsequent reflective process is forgotten in many cases.

Although no one would question the introduction of this competence in university studies, we must not ignore the challenges and difficulties inherent in assessing teamwork. Since much of the work is done outside the classroom, such assessment raises several questions: How can the professor differentiate between work done by the individual and the team?; How can evidence be gathered of the learning and contributions each student makes in a working group?; How do we assign the right mark to those students who have contributed little and taken advantage of their classmates' work?

Peer assessment can be considered a support tool for professors in assessing teamwork (Sergi 2007). It is the team members themselves who are best able to assess the dynamics of this work and the degree to which each of the members assimilates the specific competence, as has been previously reported by Gransberg (2010). Consequently, peer assessment can constitute an effective tool in providing information regarding both the participation of each team member in the final outcome and their abilities and shortcomings with regard to teamwork.

Several authors have made contributions and conducted studies on self-assessment and peer assessment in teamwork. In general, such studies conclude that these techniques are effective and that students' perceptions are usually favourable (Sergi 2007; Carson and Glaser 2010; Gransberg 2010; Neus 2011; Weaver and Esposto 2012). Moreover, there are numerous articles describing how different assessment tools and mechanisms are applied to teamwork. Thus, for example, with regard to differentiation between group and individual grading, Neus (2011) and Weaver and Esposto (2012) propose coassessment systems by means of which students grade their peers, with each individual mark being obtained from the average of the grades received from peers and the group mark awarded by the professor. According to these authors, systems supported by peer feedback are more accurate than those which only use feedback from the professor.

More complex studies also highlight peer assessment as a very useful tool for obtaining an individual mark that reflects the contribution and participation of each student during teamwork. Sergi (2007) presents a study on peer assessment conducted over 9 years, evaluating the participation and involvement of students in several working groups. This procedure is known as long-term peer assessment. For one semester, students participate in three different working groups and are assessed by members of the three teams regarding their contribution to and participation in groupwork, meaning each student is assessed by an average of 12 colleagues. This system brings a balance to peer assessment and minimizes the low marks that can result from a clash of personalities and aims. Jin (2012) presents a study consisting in implementing two peer assessment methods, concluding that this type of assessment may be useful to adequately measure each student's participation and degree

of commitment, as well as increasing synergy between group members. Gransberg (2010) describes a procedure for peer assessment used in 325 groups (1,200 students). The study applies an adaptation of the Croft Peer-Evaluation Method (Croft, Meyers, and Fentimen 1995). For example, if a team of three is awarded a mark out of 100, the maximum number of accumulated points would be 300. In this example, if students believe that each member has contributed equally, the accumulated points will be distributed evenly. By contrast, and to cite an extreme case, if one member did not participate in the work, the other two would share the points and the one who did not work receives a score of 0. Each student has to carry out a self-assessment and also assess the other team members. The study confirms that students take the opportunity to differentiate between individual marks within teamwork, since in half of the cases it is not an equal distribution, confirming that students who contribute more are rewarded and those who contribute less are penalized.

Finally, Anson and Goodman (2014) postulate that the key to improving teamwork is to provide quality feedback on both an individual and group level. Only then will students be able to learn how to improve their behaviour. It is based on the hypothesis that slackers are a frequent problem in teamwork, often leading to uncomfortable situations of non-communication within the group, which end up affecting the motivation of its members. The authors present four points to consider with regard to improving group dynamics: anonymity, offering groups different opportunities to reflect, posing open-ended questions to make students engage in constructive reflection, and holding group discussions for improvement after assessment.

Objectives and methodology

In order to provide more information on how the peer assessment system works with regard to teamwork, a study was undertaken during the 2012-13 academic year by the University of Girona's Network of Educational Innovation in Evaluation, which comprises professors from different disciplines and four different faculties and centres (Education, Humanities, Science and Engineering). The main objective of this study was to analyse how the peer assessment system works with regard to teamwork on degree courses in different disciplines and its degree of acceptance by students. The study also aimed to foster the participation and involvement of students in evaluating their own work and that of their peers. The assessment system applied was a modified version of the method first posited by Croft (Croft, Meyers, and Fentimen 1995).

The study involved a group activity being undertaken by students, who then assessed the teamwork that had taken place and expressed their opinion on this assessment method. It was performed in six subjects - two from Social Sciences, one from Science and three from Engineering - on five different degree courses, with the participation of 243 students (Table 1). The number of students who took part in the activity ranged from 27 to 59, divided into 9-15 groups of 3 to 5 members each. The group activity consisted in preparing and making oral presentations, and producing written assignments, projects, spreadsheets or business plans. Aside from the content involved in each activity, these educational activities also aimed to foster students' active participation in teamwork, and have them take responsibility for and involve them in assessment. To add weight to the work of student assessment, in each subject the mark awarded by peers was worth 5% to 50% of the final mark for the subject (depending on the characteristics of the subject, the activities carried out and the type of assessed activity). None of the students had previous experience in the use of peer assessment to evaluate teamwork, although some of them had experience on the use of peer assessment in other activities (oral presentations, individual essays, etc.).

Course	Project design	Tourism, territory and environment	Social and legal aspects of biotechnology	Programming project	Structures	Business organization and management
Subject	Social education	Geography, spatial planning and environmental management	Biotechnology	Computer Engineering	Agricultural Engineering	Agricultural Engineering
Field	Social Sciences	Humanities	Science	Engineering	Engineering	Engineering
Year	Second	Third	Third	Second	Second	First
Number students (Number groups)	58 (15 groups)	29 (9 groups)	59 (12 groups)	38 (14 groups)	32 (8 groups)	27 (10 groups)
Activity	Project work	Written work	Oral presentation	Programming project	Produce a spreadsheet	Group work (business plan)
% of final course mark	45%	40%	5%	50%	5%	20%
Previous experience in peer assessment activities	Yes	No	Yes	No	No	No
Number of students who responded to questionnaire	32	29	23	33	19	27
% of groups with unequal mark distribution	40	67	33	57	0	10

Table 1. Overview of the characteristics of subjects considered in the study.

Despite the heterogeneity of subjects, the same methodology was followed in all of them:

- (1) Joint preparation of documents for students to carry out self-assessment and peer assessment of teamwork by all members of the University of Girona's Network of Educational Innovation in Evaluation.
- (2) A group activity carried out in each subject.
- (3) Professor grades the activity.
- (4) Individual and group assessment of teamwork by students. To facilitate reflection on and assessment of the process and work done, a rubric and two reports were provided for evaluation: one individual and the other group. In the individual report, students assess themselves and colleagues, while in the group report; the group assesses all of the teamwork together. In both cases, evaluation is on a scale of 1 to 4 and students can provide arguments and evidence. The assessment indicators included in these two reports are related to the eight basic behavioural components reviewed by Baker (2008): attendance at group meetings, dependability and keeping his or her word, quality of work, effort and/or extra effort, cooperation and communication with other members, management of group conflict, making cognitive contributions, and structuring group work for goal achievement.
- (5) Distribution of the activity marks between members of each group based on the reflective work done in individual and group reports. For example, if the final teamwork mark is a 7 and the team is composed of four students, the group has 28 points to distribute among the members of the team. The groups have the autonomy to decide by consensus whether the mark should be the same for all members, or whether to differentiate based on the level of involvement and work done by each. Once decided, they must submit a report to the professor with the mark distribution. The mark in the report is the one given to each student as a final mark for teamwork.
- (6) Each student responds to an anonymous survey regarding their perception of the teamwork assessment (Table 2).
- (7) Gathering of marks and results of surveys in each subject.
- (8) Joint analysis and discussion of results for all subjects by all members of the Network of Educational Innovation in Evaluation.

Prior to the activity, the professors responsible for the subjects provided students with a detailed explanation of the teamwork assessment process. Although some students had experience in the use of peer assessment, they had not previously used this method to assess teamwork.

The aims of this study were twofold: to analyse student responses to questionnaires regarding their perception of the methodology used (Table 2), and to relate the results of distributing marks for the activity among different groups to students' opinions of this assessment methodology.

	1	2	3	4
1. I was able to assess my colleagues in relation to the teamwork process				
2. The professor clearly explained the procedure for carrying out an effective assessment of teamwork				
3. The professor provided me with the tools to carry out an effective assessment of teamwork				
4. Assessing teamwork meant assuming a lot of responsibility for me				
5. Knowing that my colleagues would be assessing me made me participate more in teamwork				
6. Knowing that there would be an assessment of teamwork made my colleagues participate more				
7. Assessing teamwork allowed me to accept my colleagues' contributions and criticism constructively				
8. Reflecting before distributing the marks meant individual marks for teamwork were fairer				
9. I agree with the mark I received for the teamwork activity				
10. This type of assessment has helped me develop skills that will be useful in my future career (evaluating CVs, proactive communication...)				
11. I would recommend continuing the method in these and other degree subjects				
Name two aspects you consider to be positive about this experience of assessing teamwork				
a)				
b)				
Name two difficulties or limitations that you have encountered in the process of assessing teamwork				
a)				
b)				
What would you change about the rubric?				

Table 2: Questionnaire

The perception questionnaire contained 11 items on a scale of 1 to 4 (where 1 means strongly disagree and 4 totally agree) and three open questions (where students had to note down two positive aspects, two difficulties or limitations of the assessment system and proposals for changes to the rubric). Each professor administered the questionnaires to their group on paper or online, and the results were entered in a statistical processing program.

We have carried out a detailed statistical analysis on both the data obtained from the perception questionnaire and the way the workgroups have distributed the mark among their members. The main target of such an analysis is the study of possible similarities and differences between the outcomes obtained from different student profiles coming from the classification of the students according to 3 different axis: the first one referring to the presence (yes/no) of student previous experience in peer-assessment activities, the second one referring to the weight (high/low) of the teamwork in the final mark, and the third axis referring to the belonging (yes/no) of the student to a technology degree.

The statistical study begins with a descriptive analysis of the results of the 11 items in the questionnaire (means, standard deviations, quartiles, number of answers). Next we show the results from an analysis of correlation between each pair of items in the questionnaire, in which the Pearson correlation coefficient (covariance divided by product of standard deviations) has been computed for the results of each pair of items, and correlation tests have been used to determine if the correlation values are significant. Thirdly, we present an analysis, carried out using a t-test, of significant differences in the questionnaire results according to the three previously described axis. Finally, we include a contrast of the teamgroup mark distribution among the group members against the three previously described axis; such a contrast has been made using an X-squared test.

Results

Descriptive analysis of the questionnaire results

Table 2 presents the mean, standard deviation, quartile values, number of answers registered (n) and number of students that have not answered the question (NA) for each of the 11 questions (Table 2) valued on an integer scale of 1-4. The distribution of 1-4 values is shown in Figure 1.

	mean	sd	0%	25%	50%	75%	100%	n	NA
Question 1	3.325	0.793	1	3.0	3	4	4	163	0
Question 2	3.292	0.652	2	3.0	3	4	4	130	33
Question 3	3.515	0.638	1	3.0	4	4	4	130	33

Question 4	2.896	0.887	1	2.0	3	4	4	163	0
Question 5	2.485	1.029	1	2.0	3	3	4	130	33
Question 6	2.531	0.950	1	2.0	3	3	4	130	33
Question 7	3.129	0.833	1	3.0	3	4	4	163	0
Question 8	3.043	0.974	1	3.0	3	4	4	162	1
Question 9	3.417	0.770	1	3.0	4	4	4	156	7
Question 10	2.902	0.833	1	2.5	3	3	4	163	0
Question 11	2.957	0.952	1	2.0	3	4	4	163	0

Table 3: Descriptive statistics

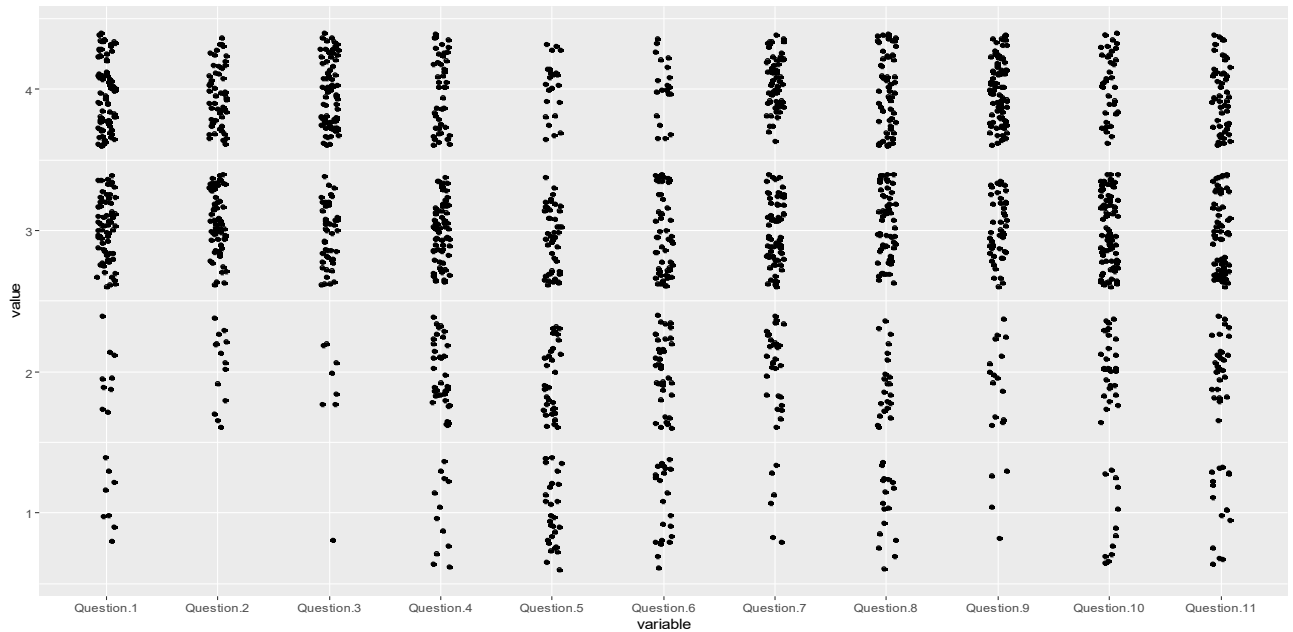


Figure 1: Distribution of values 1-4 for each of the 11 questions in the questionnaire

An analysis of responses to the questionnaires reveals that overall students had a positive opinion of peer assessment for teamwork (Figure 1). Note that the mean values range from 2.485 (Question 5) to 3.515 (Question 3) on a [1-4] scale (Table 3). The results for Question 11, which asks about the convenience of applying the method in further years, are of particular interest as they show a mean near 3, which means that the students are mostly in favour of this method.

Students' favourable opinion of the method can be explained by the fact that most professors in the Network of Educational Innovation in Evaluation have had previous experience in these activities. Students awarded high scores to professors' ability to clearly explain the procedure (Question 2) and provide them with sufficient tools (Question 3) for them to perform an effective assessment of teamwork. The high average and low variability score across all subjects for these two questions (Figure 1) is representative of this favourable opinion.

Questions 5 and 6, related to students' perception of whether peer assessment increased their involvement or that of their peers in groupwork, were the ones that received the lowest average score, about 2.5, but with a great variability. In this respect, students recognized that they did not involve themselves or cooperate more in teamwork purely because they knew they would be assessed by their peers.

In general terms, students tend to consider that this type of assessment can help them develop skills that will be useful in their future careers. This was addressed in question 10, which obtained a mean score close to 3.

Correlation analysis

A correlation analysis was performed for the answers to each pair of items in the questionnaire. Correlations were computed, and their significance (or lack thereof) determined by means of a correlation test. The aim of such an analysis is to detect pairs of questions in which the correspondence between students' answers is significantly high and draw conclusions from this. Table 4 shows the correlation values corresponding to the results of each pair of questions over the diagonal, and value p from the correlation test under the diagonal, which is used to determine the significance of the corresponding correlation score. Values of p lower or equal to 0.05 (in bold) mark the correlation scores assumed not to be 0.

	Question 1	Question 2	Question 3	Question 4	Question 5	Question 6	Question 7	Question 8	Question 9	Question 10	Question 11
Question 1		0.220	0.255	-0.024	0.039	0.102	0.278	0.388	0.252	0.140	0.126
Question 2	0.012		0.510	0.020	0.159	0.184	0.142	0.270	0.280	0.338	0.286
Question 3	0.004	<.001		0.140	0.199	0.156	0.279	0.345	0.354	0.353	0.251
Question 4	0.790	0.818	0.114		0.311	0.361	0.244	0.133	0.141	0.264	0.194
Question 5	0.661	0.072	0.024	0.001		0.746	0.493	0.283	0.000	0.285	0.180
Question 6	0.252	0.037	0.077	<.001	<.001		0.457	0.297	0.074	0.238	0.194
Question 7	0.001	0.109	0.001	0.005	<.001	<.001		0.512	0.270	0.438	0.410
Question 8	<.001	0.002	<.001	0.132	0.001	0.001	<.001		0.477	0.351	0.565
Question 9	0.004	0.001	<.001	0.112	0.999	0.404	0.002	<.001		0.383	0.493
Question 10	0.112	<.001	<.001	0.002	0.001	0.007	<.001	<.001	<.001		0.559
Question 11	0.156	0.001	0.004	0.028	0.041	0.027	<.001	<.001	<.001	<.001	

Table 4: Correlation values (over the diagonal) and p values (from the correlation test, under the diagonal) corresponding to each pair of questions.

It can be observed that although all the correlation values are positive, not all of them are significant according to the correlation test. We would emphasize the following pairs of questions, as their significant correlation value is especially meaningful:

- Question 1 (on student ability to peer-assess) and Question 8 (on fairness of individual marks obtained).
- Question 2 (on clear procedure explanation by the professor) and Question 11 (on favourable opinion of the procedure and recommending its continuation).
- Question 9 (on student agreement with her own mark) and Question 11 (on favourable opinion of the procedure and recommending its continuation).

Moreover, there seem to be intercorrelation patterns among:

- Questions 1, 2, and 3 (which refer to perception of students' willingness to evaluate their peers, clear explanation of the method by the professor, and provision of evaluation tools by the professor)
- Questions 4, 5, and 6 (which refer to perception of the increase in students' own responsibility and that of peers)
- Questions 7 and 9 (which refer to perception regarding acceptance of peer criticism and fairness of the mark received)

Analysis of differences according to previous experience in peer-assessment, weight of teamwork in the final mark, and technology or non-technology degree

We have analysed possible differences in answers to the issues on the questionnaire according to three axes: previous experience in peer-assessment, weight of teamwork in the final mark and technology or non-technology degree. For all three axes, the relative number of observations in each category is sufficient to guarantee the validity of the results obtained:

- 55 of the 163 students in our study had previous experience of peer-assessment activities, which corresponds to around 1/3 of the total.
- 69 and 94 (of 163) students were involved in subjects where the weight of teamwork is low and high, respectively.
- 79 of the 163 students were studying a technology degree and the rest (84) a non-technology degree.

With regard to students' previous experience, according to the t-test based contrast (Table 5), the results of all questions – except for 2, 4, 5, and 6 – show significant differences (i.e. a p-value under 0.05). The significance of the difference is especially noteworthy in two of the questions: for Question 11, referring to continuation of the method in the same and other subjects, we see that students with previous experience in peer-assessment activities are much more in favour of continuation of the method than students with no previous experience; and in Question 7, we clearly observe that students with previous experience seem to be more willing to accept both criticism and contributions from their team peers.

Previous experience	NO (108)		YES (55)		t statistic	p-value
	mean	sd	mean	sd		
Question 1	3.231	0.849	3.509	0.635	-2.346	0.0204
Question 2	3.253	0.699	3.345	0.584	-0.817	0.4157
Question 3	3.387	0.695	3.691	0.505	-2.891	0.0045
Question 4	2.880	0.894	2.927	0.879	-0.325	0.7455
Question 5	2.467	1.070	2.509	0.979	-0.235	0.8148
Question 6	2.613	0.971	2.418	0.917	1.169	0.2446
Question 7	2.972	0.859	3.436	0.688	-3.737	0.0003
Question 8	2.907	0.996	3.309	0.879	-2.636	0.0095
Question 9	3.304	0.818	3.630	0.623	-2.777	0.0063
Question 10	2.815	0.877	3.073	0.716	-2.011	0.0464
Question 11	2.676	0.965	3.509	0.635	-6.598	<0.0001

Table 5: Differentiation according to students' previous experience.

The t-test based contrast (Table 6) for the weight of peer-assessed teamwork in the final mark reveals that only Questions 1, 5, 7, and 8 (in bold) show significant differences in the final mark for the subject (p-value under 0.05). Thus, on the one hand, peer-assessment makes students whose teamwork has a high weight in the final mark feel more involved in the teamwork (Question 5), accept criticism from their group mates in a more constructive way (Question 7), and consider their individual mark in the teamwork to be fairer (Question 8); and on the other hand, with such high-weighted teamwork students seem to feel slightly less capable of evaluating their group mates (Question 1).

Weight percent	LOW 5-20% (69)		HIGH 40-50% (94)		t statistic	p-value
	mean	sd	mean	sd		
Question 1	3.464	0.655	3.223	0.869	2.013	0.0458
Question 2	3.304	0.734	3.279	0.552	0.2269	0.8209
Question 3	3.464	0.719	3.574	0.531	-0.9994	0.3196

Question 4	2.783	0.953	2.979	0.830	-1.3706	0.1728
Question 5	2.275	1.042	2.721	0.968	-2.529	0.0127
Question 6	2.420	0.961	2.656	0.929	-1.419	0.1584
Question 7	2.971	0.939	3.245	0.729	-2.0162	0.0459
Question 8	2.841	1.066	3.194	0.875	-2.2457	0.0264
Question 9	3.449	0.738	3.391	0.797	0.4740	0.6362
Question 10	2.855	0.928	2.936	0.759	-0.5944	0.5533
Question 11	2.811	1.019	3.064	0.890	-1.6469	0.1019

Table 6: Differentiation according to weight of teamwork in the final mark.

In relation to doing a technology or non-technology degree, the t-test based contrast (Table 7) shows significant differences in the answers of questions 3, 5, 6, 7, 8, and 11. Compared to students from non-technological degrees, students from technological ones (i) evaluate worse the professor explanation of the assessment method (Question 3), (ii) feel less that own (Question 5) and groupmate's (Question 6) motivation increase due to the peer-assessment method, (iii) tend to be less receptive to teammates' contribution/criticisms (Question 7), (iv) consider less important the contribution of the peer-assessment method to the fairness of the individual marks (Question 8), and finally (v) are less in favour of the continuation of the method in the same or other subjects (Question 11). In sum, students with a technological profile seem to be less enthusiastic about this peer-assessment method than non-tech students are.

Moreover, open questions in the survey show a somewhat unfavourable opinion of the peer-assessment method used among students on technology degrees. Their argument was based on the responsibility for assessment falling on the professor and the fact that peer assessment may create problems between students in the classroom. Two responses to such open questions were:

- "Possible conflicts / tensions between classmates (distribution of marks)".
- "It may lead to confrontations".

(Agricultural Engineering Degree Students)

Technology	NO (88)		YES (75)		t statistic	p-value
	mean	sd	mean	sd		
Question 1	3.440	0.717	3.203	0.853	1.921	0.057
Question 2	3.369	0.576	3.152	0.759	1.689	0.095
Question 3	3.619	0.536	3.326	0.762	2.314	0.024
Question 4	2.964	0.842	2.823	0.930	1.016	0.311
Question 5	2.702	0.967	2.087	1.029	3.331	0.001
Question 6	2.679	0.907	2.261	0.976	3.936	0.019
Question 7	3.369	0.690	2.873	0.897	3.936	<0.001
Question 8	3.238	0.887	2.833	1.025	2.680	0.008

Question 9	3.446	0.800	3.384	0.738	0.505	0.614
Question 10	2.976	0.776	2.823	0.888	1.172	0.243
Question 11	3.274	0.797	2.620	0.991	4.622	<0.001

Table 7: Differentiation according to studying a technology degree.

The relationship between mark distribution and student type

We carried out a statistical analysis of the relationship between the marks distributed among group members and student type in accordance with the 3 axes mentioned above (students' previous experience of peer-assessment, weight of teamwork in the final mark, technology or non-technology degree), and analyzed students' individual and group reports on teamwork.

In the statistical analysis, a ratio contrast based on a distribution approximated to a Gaussian was carried out using a X-squared with 1 degree of freedom, resulting in X-squared values, which correspond to p-values greater than 0.05 (Table 8) in Axes 1 and 3, and lower than 0.05 in Axis 2. Thus, we conclude that there are no significant differences in mark distribution in relation to either students' previous experience or doing a technology degree, while the weight of the teamwork mark is significant. This indicates a greater tendency towards an unequal distribution within the group when teamwork is awarded a greater weight in the final mark. This was found to be the case in three degree subjects (Table 1), where the weight of the mark is between 40% and 50% and the unequal distribution of marks among group members is between 40% and 67%.

	EVEN	UNEVEN	NUMBER OF GROUPS	X-SQUARED VALUE	p-VALUE
Previous experience					
NO	26 (63.4%)	15 (36.6%)	41	0.0014285	0.9699
YES	17 (63.0%)	10 (37.0%)	27		
Mark weight					
HIGH (40-50%)	18 (47.4%)	20 (52.6%)	38	9.3275	0.002
LOW (5-20%)	25 (83.3%)	5 (16.7%)	30		
Technology degree					
NO	20 (55.6%)	16 (44.4%)	36	1.941	0.164
YES	23 (71.9%)	9 (28.1%)	32		

Table 8. Relationship between mark distribution (even or uneven) and students' previous experience, weight of teamwork in the final mark and doing a technology degree

Among the technology degrees, there were two subjects with a greater proportion of groups choosing to distribute the teamwork mark equally (Structures 100% of groups, and Business organization and management 90%, Table 1). These were students on first and second-year courses who did a groupwork evaluation activity that has little weight in the final grade for the subject. Therefore, there was low competition and great solidarity among them, as well as less maturity and critical ability (as seen from the reports). These students also had a less favourable opinion of peer assessment and scored lowest on the survey questions (Figure 2).

The analysis of the individual and group reports reflecting on groupwork revealed that an equal distribution is indicative of the group having functioned properly. In the case of groups that distributed marks unevenly, it was observed that some groups chose this unequal distribution to penalize members who had not worked enough, while other groups rewarded students who deserved a higher mark for their greater contribution or for having acted as leaders.

Course	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11
Project design	3,44	3,16	3,66	3,09	2,41	2,19	3,38	3,31	3,58	3,16	3,59
Tourism, territory and environment	3,31	3,41	3,48	3,03	3,07	3,17	3,24	3,10	3,10	2,79	2,83
Social and legal aspects of biotechnology	3,61	3,61	3,74	2,70	2,65	2,74	3,52	3,30	3,70	2,96	3,39
Programming project	2,94			2,82			3,12	3,16	3,48	2,85	2,76
Structures	3,42	2,84	3,11	2,37	2,00	2,21	2,63	2,32	3,16	2,53	2,37
Business organization and management	3,37	3,37	3,48	3,15	2,15	2,30	2,74	2,81	3,44	3,00	2,63
Average	3,35	3,28	3,49	2,86	2,46	2,52	3,11	3,00	3,41	2,88	2,93

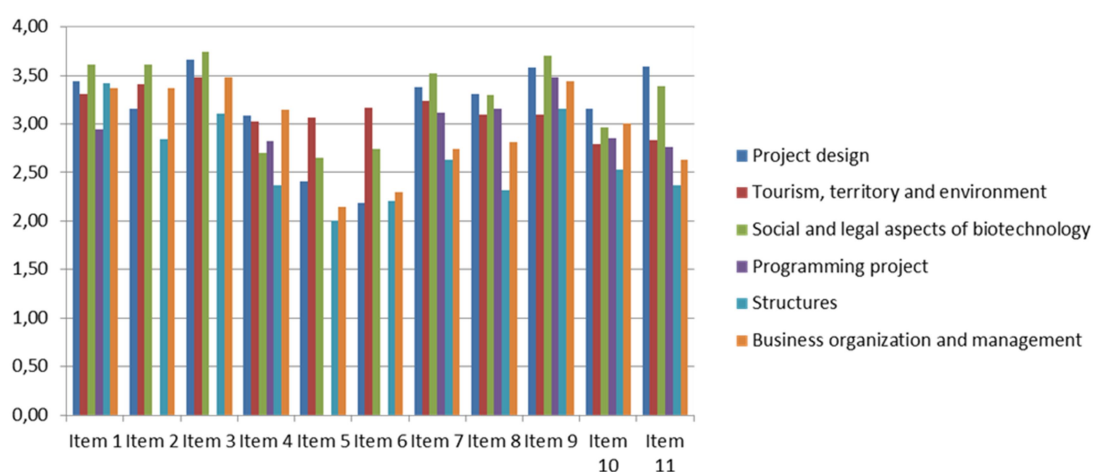


Figure 2: Average results for each item

Discussion and conclusions

The inclusion of the competence of teamwork on the curricula of many degree courses in the Spanish university system has prompted a debate on how to coherently assess this competence for each member of the team. The main problem faced by professors is detecting and differentiating between those students who have participated most and least in teamwork (Fellenz 2006; Carson and Glaser 2010; Gransberg 2010; Anson and Goodman 2014). The present study demonstrates that peer assessment is useful in reducing this difficulty associated with groupwork. Students assessed themselves and other members of their team using the methodology for assessing teamwork presented in this article, with the mark assigned by the professor being distributed among group members. As expected, this system allowed for a distinction to be made between students' varying degrees of involvement and contribution. In general, we observed less balanced distributions in the subjects where the teamwork mark carried an important weight in the final grade.

Students had full autonomy and freedom to distribute the mark for the activity as they wished. Some authors believe that this autonomy leads to all group members agreeing to give themselves the same mark and recommend taking steps to prevent this (Abelson and Babcock 1985). In contrast, other authors consider that to ensure a fair distribution of marks, students should feel free to distribute them as they see fit (Fellenz 2006; Gransberg 2010). Our study reinforces the position of the latter authors given that even though there was no requirement for differential evaluations some groups still opted for an unequal distribution of the mark.

Our analysis of the survey responses revealed some aspects common to all subjects. Firstly, this study has shown that, overall, the assessment methodology applied is well accepted by students. Positive attitudes towards the use of peer assessment activities have also been previously reported (Fellenz 2006; Wen and Tsai 2006; Sergi 2007; Carson and Glaser 2010; Gransberg 2010; Neus 2011; Weaver and Esposto 2012), and student attitudes towards groupwork have even been described as more positive when peer assessment is used (Chapman and van Auken 2001). This positive perception can be explained by the advantages of peer assessment compared to the traditional system of evaluation, which assigns a mark to all members of a workgroup. On the one hand, the process of self and peer-assessment in teamwork helps students to become aware of both the process and the importance of working together effectively. Moreover, this method has a formative character that allows students to obtain feedback on how to develop the competence of teamwork and find out about their strengths and weaknesses, as Sergi (2007) and Anson and Goodman (2014) also point

out. At the same time, peer assessment also allows students to acquire experience in the process of evaluating and judging the learning of others. In short, this evaluation system improves the quality of students' performance (Fellenz 2006) and leads to an improved student experience (Pope 2005).

The second common aspect observed refers to students' involvement in groupwork. Although it has been noted that this evaluation system encourages students to work harder (Pope 2005), become more involved in the team (Jin 2012) and accept classmates' criticism constructively, this was not the perception of the majority of participants in our study.

Moreover, in accordance with previous reports (Hanrahan and Isaacs 2001), overall, all participants in the study believed that the peer assessment activity could help them to develop skills that will be useful for their career. In fact, teamwork is one of the skills most valued by recruiters, even if both graduates and employers agree that students finishing their degree have not developed this competence to any great extent.

Despite the above generalizations, differences were observed in the perceptions of different types of students regarding this assessment methodology for teamwork. Previous studies have highlighted this variability in students' perceptions regarding use of the methodology. As noted by Fellenz (2006), this diversity in perceptions ranges from studies finding great enthusiasm among students to others in which it is considered a fair and equitable method, and then those in which it is considered inefficient for solving the free-rider problem. The study presented here allows us to explain some of the causes of this diversity of perceptions.

Our analysis of the responses to the survey show significant differences according to the weight of peer-assessed teamwork in the final mark awarded for the subject. The more weight awarded to the mark, the more involved students seem to be in teamwork, the more they accept criticism from their group mates in a more constructive way, and the fairer they consider the mark obtained for teamwork, although it is also worth noting the less they consider themselves capable of evaluating their peers. Students with previous experience seem to be more willing to accept both the criticism and contributions of their team peers, and had a much more favourable opinion. Thus, this student profile is much more in favour of the method being applied in other subjects and degree courses. The fact that students with prior experience have a more favourable opinion than those without experience, it had been reported previously (Wen and Tsai 2006). In addition, although some students quickly accept this kind of experience, we know that others express difficulty in implementing this type of process and therefore lack confidence in it (Fellenz 2006). These students consider assessment to be the responsibility of the professor and want to keep their distance from it.

Moreover, the favourable opinion of students with experience in peer assessment shows that if students engage in such activities, they adapt to peer assessment experiences and take on more responsibility for them. These results suggest that participation in peer assessment activities helps to reduce students' reluctance to participate in them.

The results also show that our professors' and students' previous experience in participatory evaluation processes was required to achieve good results. Furthermore, we have found that it is very important to explain the assessment process well and provide students with rubrics and reflection sheets that help with this process and facilitate understanding of these tools (Planas et al. 2014). In agreement with previous experiences, this explanation should come before starting the peer assessment activity and include organized and well-defined guidelines (Falchikov and Goldfinch 2000; Baker 2008). Professors and students can have different conceptions not only of the objectives of an activity, but also of the evaluation criteria (Norton 1990). A detailed explanation of the assessment system allows students to understand the learning objectives, so that they are then able to apply the assessment criteria and assess their peers more correctly (Black and William 1998; Falchikov and Goldfinch 2000; Lane 2007; Cestone, Levine and Lane 2008). It is possible to thus increase students' acceptance of this assessment method, especially in the case of groups with no previous experience of peer evaluation (Fellenz 2006). This also helps students to perceive peer assessment as a serious process and show them what is expected of them with regard to groupwork. In fact, the opinions of our students confirm the importance of the explanation prior to starting the process. As reported by Wen and Tsai (2006), students get frustrated when they are not provided with clear guidelines and references.

When planning this research, the professors involved thought it fit to include students' reflections as part of the activity. Requiring students to give reasons to justify their marks and the final mark distribution means the individual and group mark obtained is the fruit of reflection, and therefore fairer. Offering students an individual and group reflection process also helps improve group performance and increases learning among its members (Anson and Goodman 2014), while also providing information regarding their strengths and weaknesses. In addition, in their reports students had the opportunity to make comments to professors and communicate the problems they or the group had faced.

Taking all the results as a whole, similarly to other research studies (Baker 2008) we have shown that peer evaluation can provide valuable information about group member performance. The peer evaluation system applied in this study accomplishes the features necessary for a useful peer

assessment method (Fellenz 2006; Baker 2008; Gransberg 2010): it is easy to implement, it is practical and easy to use for students and professors, it allows the contribution of each member in the teamwork to be scored following an easy process, and, at the same time it motivates high levels of positive behaviours and is perceived as fair, tackling free-riding and intragroup conflict. This evaluation system can also be applied to different activities and in different disciplines.

At Spanish universities very few formative experiences have been carried out that offer students reflective space to analyse the dynamics developed in teamwork or the personal contributions and roles developed by each member. To ensure that university graduates become professionals skilled in teamwork there is a need to introduce reflexive evaluative spaces, methodologies and processes that allow students to analyse their own performances and those of their peers. Unfortunately, Spanish higher education has yet to evolve much in this direction. Currently, even though there is a tendency to award more importance to teamwork in general and group assessment experiences in particular, university professors award much more importance to the vertical flow between professor and student than the horizontal interaction among students.

Despite many Spanish universities being relatively young, as is the case with the University of Girona, they have innovated very little in evaluation methods. Students mostly continue to be evaluated on the basis of knowledge and using traditional methods. However, when evaluating teamwork many professors are concerned to ensure that students receive a fair mark in terms of the skills they have acquired and their level of involvement. The results of this study demonstrate that self-evaluation and peer evaluation are good alternatives to conventional methods, as evidenced by the fact that students who have experience in them rate these methods highly and believe they should be extended to more subjects and degree courses. We believe it is worth conducting further research in this field and disseminating the lessons learned in order to encourage more professors to apply these and other new evaluation methods.

References

- Abelson, M. A., and Babcock, J. A. (1985). Peer evaluation within group projects: A suggested mechanism and process. *Organizational Behavior Teaching Review*, 10(4): 98-100.
- Anson, Rl, and Goodman, J. 2014. "A Peer Assessment System to Improve Student Team Experiences.", *Journal of Education for Business* 89 (1): 27-34. Doi: 10.1080/08832323.2012.754735

- Baker, D. F. 2008. "Peer assessment in small groups: A comparison of methods". *Journal of Management Education*, 32(2): 183–209. Doi: 0.1177/1052562907310489
- Black, P., and Wiliam, D. 1998. "Assessment and Classroom Learning." *Assessment in Education* 5 (1): 7–74. Doi: 10.1080/0969595980050102
- Carson, K. M., and Glaser, R. E. 2010. "Chemistry Is in the News: assessing intra-group peer review." *Assessment & Evaluation in Higher Education* 35(4): 381-402. Doi: 10.1080/02602930902862826
- Cestone, C. M., Levine, R. E., and Lane, D. R. 2008. "Peer-assessment and evaluation in team-based learning". *New Directions for Teaching and Learning* 116: 69-78. Doi: 10.1002/tl.334
- Chapman, K. J., and van Auken, S. 2001. "Creating positive group project experiences: An examination of the role of the instructor on students' perceptions of group projects". *Journal of Marketing Education* 23 (2): 117-127. Doi: 10.1177/0273475301232005
- Croft, F. M., Meyers, F. D., and Fentimen, A. W. 1995. "An algorithm for evaluating team Projects". *Engineering Design Graphics Journal*, 59(3): 18–20.
- Falchikov, N., and J. Goldfinch. 2000. "Student Peer-assessment in Higher Education: A Meta-analysis Comparing Peer and Teacher Marks." *Review of Educational Research* 70 (3): 287–322. Doi: 10.3102/00346543070003287
- Fellenz, M.R. 2006. "Toward fairness in assessing student groupwork: a protocol for peer evaluation of individual contributions." *Journal of Management Education* 30 (4): 570-591. Doi: 10.1177/1052562906286713
- Gransberg, D. 2010. "Quantifying the Impact of Peer Evaluations on Student Team Project Grading." *International Journal of Construction Education and Research* 6 (1): 3-17. Doi: 10.1080/15578771003590326
- Hanrahan, S. J, and Isaacs, G. 2001. "Assessing self- and peer assessment: the students' views." *Higher Education Research and Development*, 20 (1): 53-70. Doi: 10.1080/07294360123776
- Jin, X. 2012. "A comparative study of effectiveness of peer assessment of individuals' contributions to group projects in undergraduate constructions management core units." *Assessment & Evaluation in Higher Education* 37 (5): 577-589. Doi: 10.1080/02602938.2011.557147

- Kearney, S. 2013. "Improving engagement: the use of 'Authentic self- and peer-assessment for learning' to enhance the student learning experience." *Assessment & Evaluation in Higher Education* 38 (7): 875-891. Doi: 10.1080/02602938.2012.751963
- Lane, D. 2007. *Engineering Feedback: A Student-developed Approach to the Assessment of Peer Evaluation in Civil Engineering*. Chicago: National Communication Association.
- Lindblom-Ylänne, S.; Pihlajamäki, H., and Kotkas, T. 2006. "Self-, peer- and teacher-assessment of student essays". *Active learning in Higher Education* 7 (1): 51-62. Doi: 10.1177/1469787406061148
- Neus, J. L. 2011. "Peer assessment accounting for student agreement." *Assessment & Evaluation in Higher Education* 36(3): 301-314. Doi: 10.1080/02602930903342315
- Norton, Linda S. 1990. "Essay-writing: what really counts?" *Higher Education* 20 (4): 411-442. Doi: 10.1007/BF00136221
- Pallisera, M., Fullana, J., Guiu, E., Planas, A., Serra, C., Soler, P., Tesouro, M. 2009. *La tutoria a la Universitat. Una proposta de Pla d'Acció Tutorial* (Dos punts, 5). Girona: Universitat de Girona
- Planas Lladó, A. ; Feliu Soley, L.; Fraguell Sansbelló, R.M; Arbat Pujolras, G; Pujol Planella, J.; Roura-Pascual, N.; Suñol Martínez, J.J; Montoro Moreno, L. 2014. "Student perceptions of peer assessment: an interdisciplinary study." *Assessment & Evaluation in Higher Education* 39 (5): 592-610. Doi: 10.1080/02602938.2013.860077
- Pope, N. K. L. 2005. "The impact of stress in self- and peer assessment." *Assessment & Evaluation in Higher Education*, 30 (1): 51-63. Doi: 10.1080/0260293042003243896
- Pujolàs, P. 2008. "Cooperar per aprendre i aprendre a cooperar: el treball en equips cooperatius com a recurs i com a contingut." *Suports* 12 (1), 21-37.
- Rodríguez, G.; Ibarra, M.S.; Garcia, E. 2013. "Autoevaluación, evaluación entre iguales y coevaluación: conceptualización y práctica en las universidades españolas." *Revista de Investigación en Educación* 11 (2): 198-210
- Sergi, M. 2007. "Evaluating short-term and long-term peer assessment of Student Teamwork." *e-Journal of Business Education & Scholarship of Teaching* 1 (1): 41-58.
- Topping, K. 2009. "Peer-assessment". *Theory into Practice* 48(1): 20-27.
- Topping, K. J., Smith, E. F., Swanson, I., and Elliot, A. 2000. "Formative peer-assessment of academic writing between postgraduate students". *Assessment & Evaluation in Higher Education* 25(2): 149–170. Doi: 10.1080/713611428

- Vickerman, P. 2009. “Student perspectives on formative peer-assessment: an attempt to deepen learning?”. *Assessment & Evaluation in Higher Education* 34 (2): 221-230.
- Weaver, D. and Esposito, E. (2012) “Peer assessment as a method of improving student engagement”, *Assessment & Evaluation in Higher Education*, 37(7): 805-816. Doi: 10.1080/02602938.2011.576309
- Wen, M.L., and Tsai, Ch-Ch. 2006. “University students’ perception of and attitudes toward (online) peer-assessment”. *Higher Education* 51(1): 27-44. Doi: 10.1007/s10734-004-6375-8