Peer and team assessment: strategies and applications in Engineering courses

Walter Aoiama Nagai[1], Rui M. Lima[2], Diana Mesquita[2], Anderson Assis Morais[1]

[1] Federal University of Itajuba, Itabira Campus, Brazil
[2] Department of Production and Systems, Centro ALGORITMI, School of Engineering, University of Minho, Guimarães, Portugal

Email: walternagai@unifei.edu.br, rml@dps.uminho.pt, diana@dps.uminho.pt, andersondeassis@unifei.edu.br

Abstract

Teamwork is one of the competencies that more often are referred has required to professional practice in Engineering. Working in teams in the learning process has been referred to be an effective way to promote the development of technical competencies while promoting the development of teamwork competencies. The students identify teamwork as motivating for their self-learning. In a teamwork environment student can deal with self-knowledge, critical analysis, knowledge of the others, individual and group performances, feedback, resilience, synergy, decision making, commitment, participation, self-esteem, leadership, and entrepreneurship. All these characteristics come from the understanding that a team is formed by individuals with different experiences, origins and individual profiles. But what are the criteria for peer or teamwork assessment? Which methods give fair rewards for different contributions to the team and its peers? There are several peer assessments studies where many experiences are described, but there are not many studies that compare the strategies of peer assessments between them. For example, when the criteria have different weights or when the scores given by the peers are anonymous or when the scores are decided by the group, and so on. The objective of this work is to describe several strategies of peer and team assessment, considering the categorization and organization carried out in order to assumptions and/or purposes of each strategy. Thus, a contribution will be made for increasing peer and teamwork assessment in Engineering courses.

Keywords: Engineering Education, Peer assessment, Teamwork, Transversal Competences, Project-Based Learning

1   Introduction

In nowadays, engineers face challenges that require a solid foundation in engineering, but also skills such as teamwork, project management, interdisciplinary problem-solving and oral and written communication (Fernandes et al., 2009; Vicente, Romano, Sá, & Lima, 2014). This interest is based on the need to improve the education of engineers to create the need to change Engineering Education (EE). New approaches should bring joy to students and teachers, and trust is the main ingredient for enjoying teaching and learning, and also for creating an atmosphere of empathy and openness that is necessary for collaboration. In order to create innovative approaches, the stakeholders need a trusting environment that will always accompany a changing environment. This environment should offer a modern, reliable, innovative atmosphere with fair assessment techniques and strategies. Besides, that classroom should contribute for the development of transversal competences like teamwork, communication, self-knowledge, critical analysis, knowledge of the others, and so on.

Project Based Learning (PBL) has been an innovative strategy in curriculum design in many engineering courses around the world. PBL is based on some fundamental principles (Graaff & Kolmos, 2003): Principle #1 - problem with a high relationship with the professional context of engineering that relevant and very significant to students; Principle #2 - interdisciplinarity that permits the students to link theory and practice and establish between different areas of knowledge; Principle #3 - teamwork with students and teachers engaged with the team learning. PBL is frequently accompanied by one or more assessment strategies of the team participants, since individual technical tests may not measure the entire learning developed by the team.

A modern engineer should know to work within interdisciplinary e multicultural teams and know to recognize the teammates' potentials for the development of the projects. The technical competences are very important, but it is very important that the engineer become a self-learner and be able to do self-monitoring to adapt to fast changing needs.

Assessment is an important element of the curriculum processes, which allow to assess the competences an engineer should develop. The traditional approach is mainly focused in knowledge assessment and technical competences, which does not correspond to the expectations of a modern engineer as referred previously. Thus, the combination of PBL and innovative strategies of assessment, e.g. peer assessment, is a way of improving engineering courses.

There are many peer assessment strategies and dimensions, which will be summarized in the next section. Nevertheless, there still a need to answer to the question: how to develop peer or teammates assessment in engineering courses? Thus, this paper will be focused in answering to this question, from the point of view of the methods identified in the literature. Particularly, this work aims to clarify and peer assessment strategies and develop the calculation details of two different methods. This paper starts by presenting a summary of classifications of peer assessment, goals, reliability and validity of existing methods, will continue by discussing grading issues and make some recommendations for applying peer assessment methods. Finally, will present some details of calculation to help teachers of engineering to apply peer assessment in their courses.

2   Peer Assessment Principles

Assessment can be seen as the determination of the amount, level, value or worth of something (Topping, 2003). But the term Peer Assessment (PA) has many meanings and terms as peer marking, peer correction, peer rating, peer feedback, peer review, and peer appraisal are some examples. It is difficult to establish relationships between assessments dimensions and students' learning due to the different purposes, functions, required levels of learning, and technical and transversals competences in engineering courses. A short guideline of peer assessment in engineering courses could be applied in different purposes and situations, that is the main subject of this work.

As described by Topping (2003), Self-Assessment (SA) is the arrangement for learners and/or workers to consider and specify the level, value or quality of their own products of performances; Boud, Cohen & Sampson (1999) and Boud & Falchikov (1989) refer to the involvement of learners in making judgements about their own learning, particularly about their achievements and the outcomes of their learning.

Team Assessment (TA) refers to individual assessments inside teams, to ensure that students have a fair recognition of their contribution to team outcomes (Brown, 1995). In the teammates assessment, the students are assessing concerning with transversal competences, because teamwork can deal with self-knowledge, critical analysis, knowledge of the others, individual and group performances, feedback, resilience, synergy, concluding requirements and deadlines, self-esteem, leadership, and entrepreneurship (Fernandes et al., 2009; Vicente et al., 2014).

Peer Assessment can be seen as an arrangement in which individuals consider the amount, level, value, worth, quality, or success of the products or outcomes of learning of peers of similar status (Gielen, Dochy, & Onghena, 2011; Gielen, Dochy, Onghena, Struyven, & Smeets, 2011; Gielen, Peeters, Dochy, Onghena, & Struyven, 2010; Harris & Brown, 2013; Pope, 2005; Topping, 2003, 2005, 2009). Peer assessment takes the form of feedback, face-to-face or otherwise, often reciprocally among the assessors and assesses (Topping, 2009). Assessment of teammates' contributions or achievements can become complex in this context of high interaction, and diverse variables, dimensions and variations between characteristics of the team. The overriding goal of peer assessment is to provide feedback to learners (Topping, 2009). Feedback can reduce errors and have positive effects on learning when it is received thoughtfully and positively. The primary goal of feedback is to enhance the performance of the individual and/or group by identifying the discrepancies between a member's expected and actual performance, thereby giving the member the opportunity to take a corrective action (Baker, 2008).

2.1   Peer Assessment Goals

The Peer Assessment permits involve the students in activities with different goals, but it is difficult to establish phases and learning outcomes when do not know how to apply the peer assessment. Besides, the peer assessment is described in many strategies, methods or strategies depending on the goals. The work by (Gielen, Dochy, & Onghena, 2011) developed an update bibliographic review from the work by (Topping, 1998), treating it as an inventory of peer assessment diversity. In this inventory was identified five clusters, twenty variables,

and a range of variations that differ from the work by (Topping, 1998). In summary, each cluster grouped some general decisions about Peer Assessment applying. **Cluster I**: **decisions concerning the use of peer assessment**, that are related to peer assessment purpose, peer assessment objects or artifacts produced, the frequency of application and the experience of the student in peer assessment, objectives or goals of peer assessment and summative or formative peer assessment's function. **Cluster II**: **link between peer assessment and other elements in the learning environment**, that are related to degree of alignment between teammates and time available of the peer assessment, relationship with other assessments and scope of involvement the peer assessment. **Cluster III**: **interaction between peers**, that are related to output object type of peer assessment goal, flow directions of peer assessment, privacy, time and place where the peer assessment occurs, role of assess between peers. **Cluster IV**: **composition of assessment groups**, that are related to how to match the students and the organization of assessors/assesses. **Cluster V**: **management of the assessment procedure**, that are related to how do peers assess, the compulsory characteristic of participation of peers, the needs of reward by participation of peers, the need by training or guidance in peer assessment, and the last variable is the quality control that relates to reactive or proactive control of learning of peers.

Specifically, the goals of decisions concerning the use of Peer Assessment are described in Cluster I and there is a need to make a relationship with one quality concept of learning presented by the peers. In (Gielen, Dochy, Onghena, et al., 2011), there are five goals, and a set of quality concepts related to them can be viewed in Figure . This figure illustrates the relation between the five goals, sub-goals of PA, and the six quality concepts proposed by Gielen, Dochy, Onghena, et al. (2011).
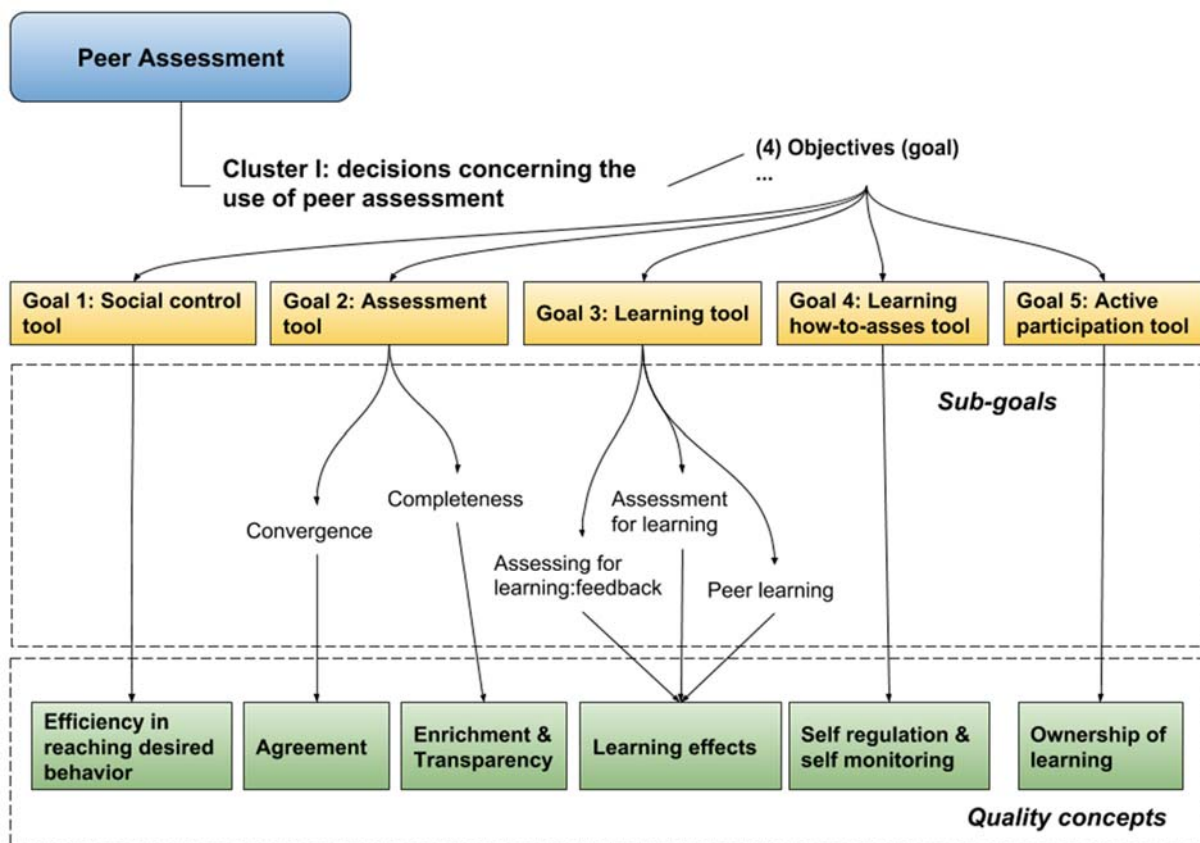

Figure 1. Illustration of the relation between goals of PA and quality concepts.

A short description of each of the five Peer Assessment goals and the quality control dimensions presented by Gielen, Dochy, Onghena, et al. (2011) are summarized in the following sentences. (1) **Social control tool**: this first goal is to make the peers assess each other using a problem's sheet with criteria to spend time on one or more tasks, and to use feedback as a social dimension with peers. Quality concept of peer assessment as social control tool: to keep the desired behavior like a controlled class. (2) **Assessment tool**: this second goal, there

are two quality concepts as convergence and completeness. The convergence quality concept is related to the agreement between students and teachers with judgments of teammates with grades and the assessments received between them. Should exist comparisons with the teacher assessment realized, teammates assessment, other episodes of assessment during the course, and the self-assessment between teammates. The completeness quality concept is related to all different opinions that enrichment for the final assessment because they contribute to its validity, and they are transparent to all participants in the assessment process and are the underlying differences in assessment schemes becoming clear. (3) **Learning tool**: the third goal states that peer assessment could be a tool for learning through learning effects (quality concept) of the way it is implemented. Thus, the authors make a distinction on three ways (named subgoals) it can be implemented as a learning tool: a student undergoing the assessment process; or a student assessing another student; or an interaction between peers. This three subgoals may contribute for the learning process. (4) **Learning how-to-assess tool**: in this goal, the students should learn to become assessors in peer assessment. This is an important part of becoming a lifelong learner. Quality concept of peer assessment as learning how-to-assess tool is self-regulation and self-monitoring of students regarding to their own learning. (5) **Active participation tool**: in this last goal, the student become an active participant in assessment process and in this way takes ownership of the learning process. Peer assessment contributes in this way to the empowerment of the student.

2.2   Reliability, Validity, and Utility of Peer Assessment

Research findings on the reliability and validity of PA mostly emanate from studies in higher education (Topping, 2009). The reliability and validity of peer assessment tend to be at least as high, and often higher, than teacher assessments (Topping, 1998, 2003). Reliability tends to be higher in relation to: the degree of advancement in the course, the nature of the product or performance assessed, the extent to which criteria have been discussed and negotiated, the nature of assessment instrumentation, the extent to which an aggregate judgment rather than detailed components are compared, the amount of scaffolding, practice, feedback and monitoring, and the contingencies associated with the assessment outcome. Irrespective of relatively high reliability, student acceptance is variable (Topping, 2003).

The reliability and validity of PA might be compromised because social processes might influence and contaminate the process. The Peer Assessments might be partly determined by friendship bonds, animosity or other power processes, group popularity levels of individuals, the perception of criticism as socially uncomfortable or even socially rejecting and inviting reciprocation, or collusion leading to lack of differentiation. The social influences might be particularly strong with *high stakes* assessment, for which peer assessments might drift toward leniency. The reliability and validity of self-assessment tend to be a lower and more variable, while the reliability and validity of peer assessment tend to be as high or higher (Topping, 2003).

3   Peer Assessment Strategies

This section is dedicated to presenting peer assessment strategies main issues, including assessment methods, grading details and some recommendations found in the peer assessment literature.

3.1   Assessment Methods

Considering a summative function in the PA, the adoption of assessment methods needs to be moderated by two important factors: first, the assessment method must be as accurate as the traditional methods existents in higher education, and second, the method must not be prejudicial to the students themselves. The teachers have used a variety of methods to assess peer performance in small groups. According to Baker (2008), some assessment methods as rate scales, single rate, allocation of points, peer comparisons, and project diaries can be used. Peer Assessment methods are described by Baker (2008) and summarized below.

The Rating Scales method is used to assess a variety of behaviors and can provide more detailed information about the rates than other methods. In a particular type of rating scale known as Behaviorally Anchored Rating Scales (BARS), each point on the rating scale is associated with a specific observable behavior that is considered critical to the success of the team, thereby reducing ambiguity.

The Allocation of points is other common peer assessment method used in small group settings involves an allocation of points based on overall contribution to group performance. For example, the number of points to be allocated is determined by multiplying the number of team member by 10, with the stipulation that the average rating for the group total was 100%.

The Peer comparisons method compares teammates with each other, that is, each teammate identifies the team member who was most outstanding on one or more performance dimensions. The points are assigned based on the number of times a student is listed by his or her peers for each dimension.

The Project diaries method assess teammates contributions made at various stages in the group project. In various checkpoints during the semester, peers were asked to list the names of the team members who had performed those specific tasks. At the end of the semester, the instructors could count how often each student was mentioned and comparing that to the maximum number of times that a team member could be mentioned. This approach serves to clarify performance expectations, ensure accountability, and reduce the effect of memory deterioration on peer assessment.

3.2   Grading Issues

The grading issues in the summative assessment of students involve many assessment factors. Therefore, the assessment methods become rates or grades that meaning the amount level learning of students. According to the description made by Baker (2008), the rating systems seem easier than ranking systems to convert into grades because they are based on absolute as opposed to relative standards. For example, the Autorating System described in Brown (1995), uses the weighting factor for each student to multiply the team rating obtained from the teacher to obtain the final rating for an individual. This weighting factor is a ratio of the average of the assessments of the individual to the average of all assessments of all individuals. Another alternative to determine the grade of ratings is based on the idea that the students divide the individual score by the total points achieved. A key concern when using peer assessment for grading purposes is the extent to which group members accept the ratings of their peers, questioning peer bias. When peer feedback affects the grade, student concerns about the grading accuracy are legitimate.

3.3   Recommendations

In the literature there are some recommendations to rating process discussed previously that are related to students' behavior in their attitudes with peer assessment. The PA offers triangulation, and thus seems likely to improve the overall reliability and validity of the assessment. A peer assessor with less skill at assessment but more time in which to do it could produce an equally reliable and valid assessment. Peer feedback should be available in greater volume and with greater immediacy than teacher feedback, which compensates for a quality disadvantage (Topping, 2009).

To avoid the leniency error when peers give all their teammates the same score, Baker (2008) recommends that students should not be allowed to give everyone the same score. To maintain marks balance and reduce the effect of an unfairly high/low rating, the top/bottom rating may also be thrown out. Alternatively, the median could be used instead of the average because there is a sometimes a tendency for peer marks to bunch around the median (Topping, 2009).

Should think about penalizing teammates that fail to contribute to peer assessment because it can compromise team performance. According described by (Baker, 2008), to remind students of the importance of their rating decisions, teachers could place an honor pledge at the bottom of the assessment instrument that states, "*To the best of my recollection and ability, the above ratings accurately reflect the performance of my peers*". This pledge hopefully encourages students to take the assessment seriously and that it is important to inform students early about how they will be assessed.

4 Implementation of two Peer Assessment methods

The implementation of a Peer Assessment method requires some levels of decision, as illustrated by Figure 2. This section proposes a model of three level of decisions: (a) Selection of an assessment method; (b) Selection of a threshold or not; (c) Utilization of absolute or relative grading? Additionally, this model is applied through the details of the implementation of two peer assessment methods.
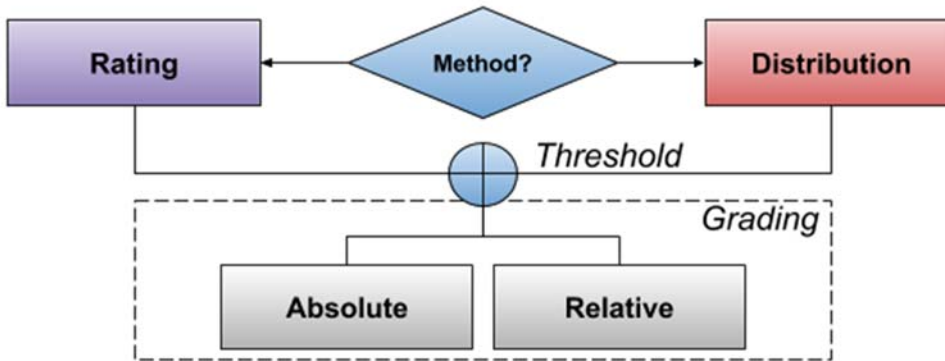

Figure 2. Three levels of decision to apply Peer Assessment Methods.

The Rating method uses a grading process is equivalent to the Rating Scale (Baker, 2008), in which students will assign a level of performance for each criteria to all peers. The Distribution method is similar to the allocation of points described in (Baker, 2008), in which students will have to decide a distribution of points among its peers. The threshold is a limit on percentage of a rating assigned/calculated in the peer assessment method. Ultimately, it is used to keep the rating received by a student between a lower and upper bound on the desired scale. Limits can be used to manage one or more students with much smaller (much larger) performance on the team. After selecting the method, it is still possible to select an absolute or relative grading process in relation to the teachers' grading. An absolute grading considers that the peer assessment will be considered as a percentage of the student's final grade. A relative grading considers the peer assessment to be used as a corrective factor to adjust the student's final grade. The next subsections will explain each level and their possible values.

4.1 Rating Method

The scale values for the Rating Method can be based on Likert scale or on other scale of values, e.g. as 1 to 10. The teacher can suggest a scale that students should agree upon. This implementation examples will consider assessment dimension as criterion defined between the students and the teacher. Some examples of criteria in PBL environments can be observed in (Mesquita, Chagas, Lima, & Chagas, 2018). For each assessment dimension, each student assigns a scale value to teammates. If there is the intention of using the self-assessment, then students should assign a value to themselves. A matrix example of one assessment dimension can be seen in Table .

Table 1. A matrix example with N peers for assessment of Dimension #1.

Dimension #1	Peer #1	Peer #2	...	Peer #N
Peer #1	$Grade_{1,1}$	$Grade_{1,2}$	...	$Grade_{1,N}$
Peer #2	$Grade_{2,1}$	$Grade_{2,2}$	...	$Grade_{2,N}$
...	...	...	...	...
Peer #N	$Grade_{N,1}$	$Grade_{N,2}$	...	$Grade_{N,N}$
	$Sum_{1,1}$	$Sum_{2,1}$	...	$Sum_{N,1}$

For each row i of Table , each student (peer i) assigns a scale value to teammates, including himself/herself. The sum ($Sum_{1,N}$) of all scale values in column j is equivalent to the total value assigned to student j, by all peers.

Table 2 presents a matrix with peer grades for M dimensions. The column *TOTAL RATE*ᵢ in each row of Table 2 is the total grade assigned to that peer by all peers and him/herself.

Table 2. A matrix example with N peers by M assessment dimensions.

	DIMENSION #1	DIMENSION #2	...	DIMENSION #M	$TOTAL\_RATE_i$
Peer #1	$Sum_{1,1}$	$Sum_{1,2}$	...	$Sum_{1,M}$	$\sum_{j=1}^{M} Sum_{1,j}$
Peer #2	$Sum_{2,1}$	$Sum_{2,2}$	...	$Sum_{2,M}$	$\sum_{j=1}^{M} Sum_{2,j}$
...	...	...	...	...	
Peer #N	$Sum_{N,1}$	$Sum_{N,2}$	...	$Sum_{N,M}$	$\sum_{j=1}^{M} Sum_{N,j}$
					$AVERAGE\_RATE$

In the last row and last column is calculated the average grade for all students denoted by *AVERAGE_RATE*, calculated using the formula of Equation 1.

$$AVERAGE\_RATE = \frac{\sum_{I=1}^{N}\sum_{j=1}^{M} Sum_{i,j}}{N * M}$$

Equation 1

The weight factor PEER_FACTOR (Equation 2) is equivalent to the ratio of each peer's TOTAL_RATE divided by AVERAGE_RATE. This PEER_FACTOR indicates each peer's assessment as seen by teammates. Instead of the average of peer's ratings, some works (Baker, 2008; Topping, 2003, 2009) refer to the use of the median of ratings. The weight factor PEER_FACTOR can be used by teachers to establish a relationship with a final grade of the team project. This relationship can be absolute or relative to the teacher's course. If a threshold is set, the weight factor can be used by the teacher to observe and motivate teammates with more difficulties in the course. The threshold can be defined in 10 to 20 percent of the PEER_FACTOR to up or down.

$$PEER\_FACTOR_i = \frac{TOTAL\_RATE_i}{AVERAGE\_RATE}$$

Equation 2

4.2 Distribution Method

The Distribution Method uses a proportional amount of points based on the number of teammates. In Baker (2008) is defined as the number of teammates multiplied by 10 but can use other value like 100, for example. Each peer should distribute all points between teammates, but the sum of all peers should not exceed the total amount of points. The idea is somehow similar to that of the Rating Method, but instead of assigning an absolute scale value to each teammate, it assigns relative values by comparison of performance. The remainder of the Distribution method calculations can be similar to those already performed using the Rating method previously described.

5 Conclusions

The objective of this study was to describe a set of peer assessment strategies and the calculation process for some methods identified in the literature. Thus, this paper make contribution in terms of what strategies has been implemented in Higher Education contexts and, particularly, how to calculate the grades of teammates. One important conclusion is that these grades should reflect the value of the contributions of each student as viewed by teammates. The grade values may use criteria or assessment dimensions agreed by students and the teacher. Given the importance of team skills in almost any organizational setting, it is good for teachers to provide group learning experiences, which can benefit from peer learning and assessment.

In (Fernandes et al., 2009; Mesquita et al., 2018) and (Kilic, 2016) are described some applying situations using peer assessments strategies with engineering course students and PBL methodology.

Some studies of peer assessment have been described and integrated aiming to help and support teachers to organize the assessment model in order to make the learning of their students more motivating. To improve the learning process and assessment for learning, decisions about peer assessment should be made intentionally, with a clear understanding of the goals of the course and the objectives of group assignments.

Acknowledgments

6   References

Baker, D. F. (2008). Peer Assessment in Small Groups: A Comparison of Methods. *Journal of Management Education*, *32*(2), 183–209. https://doi.org/10.1177/1052562907310489

Brown, R. W. (1995). Autorating: getting individual marks from team marks and enhancing teamwork. *Proceedings Frontiers in Education 1995 25th Annual Conference. Engineering Education for the 21st Century*, *2*, 3c2.15-3c2.18. https://doi.org/10.1109/FIE.1995.483140

Fernandes, S., Mesquita, D., Lima, R. M., Faria, A., Fernandes, A., & Ribeiro, M. (2009). The impact of peer assessment on teamwork and student evaluation: a case study with engineering students. *Innovation and Assessment of Engineering Curricula*. Presented at the Valladolid. Valladolid.

Gielen, S., Dochy, F., & Onghena, P. (2011). An inventory of peer assessment diversity. *Assessment & Evaluation in Higher Education*, *36*(2), 137–155. https://doi.org/10.1080/02602930903221444

Gielen, S., Dochy, F., Onghena, P., Struyven, K., & Smeets, S. (2011). Goals of peer assessment and their associated quality concepts. *Studies in Higher Education*, *36*(6), 719–735. https://doi.org/10.1080/03075071003759037

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction*, *20*(4), 304–315. https://doi.org/10.1016/j.learninstruc.2009.08.007

Graaff, E. de, & Kolmos, A. (2003). Characteristics of Problem-Based Learning. *International Journal of Engineering Education*, *19*(5), 657–662.

Harris, L. R., & Brown, G. T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: Case studies into teachers' implementation. *Teaching and Teacher Education*, *36*, 101–111. https://doi.org/10.1016/j.tate.2013.07.008

Kilic, D. (2016). An Examination of Using Self-, Peer-, and Teacher-Assessment in Higher Education: A Case Study in Teacher Education. *Higher Education Studies*, *6*(1), 136. https://doi.org/10.5539/hes.v6n1p136

Mesquita, D., Chagas, R. L. C. P., Lima, R. M., & Chagas, J. M. (2018, February). *Assessment Models in Two Project Based Learning (PBL) Approaches: an exploratory study*. 612–619. Brasilia, Brazil.

Pope, N. (2005). The impact of stress in self- and peer assessment. *Assessment & Evaluation in Higher Education*, *30*(1), 51–63. https://doi.org/10.1080/0260293042000243896

Topping, K. J. (1998). Peer Assessment between Students in Colleges and Universities. *Review of Educational Research*, *68*(3), 249–276. https://doi.org/10.2307/1170598

Topping, K. J. (2003). Self and Peer Assessment in School and University: Reliability, Validity and Utility. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising New Modes of Assessment: In Search of Qualities and Standards* (Vol. 1, pp. 55–87). https://doi.org/10.1007/0-306-48125-1_4

Topping, K. J. (2005). Trends in Peer Learning. *Educational Psychology*, *25*(6), 631–645. https://doi.org/10.1080/01443410500345172

Topping, K. J. (2009). Peer Assessment. *Theory Into Practice*, *48*(1), 20–27. https://doi.org/10.1080/00405840802577569

Vicente, S., Romano, D., Sá, V., & Lima, R. M. (2014). The Influence of Peer Assessment to the Teamwork Dynamics in Project- Based Learning. ID45.1-ID45.8.