Lack of interest? Self and peer assessment as a means to improve students' engagement

Jorge Díez

Computer Sciences Department, Escuela Politécnica de Ingeniería de Gijón, Universidad de Oviedo, Gijón, Spain

Correspondence details: Jorge Díez, jdiez@uniovi.es , Sedes Departamentales Oeste - office 1.1.32 - Campus de Gijón, 33204, Gijón, Spain

Dr Jorge Díez is an Associate Professor of the Computer Science Department in the University of Oviedo. He has written several papers about peer assessment of open-response questions and about how to use machine learning algorithms to improve the assessment. One of his interests is how to use the assessment in the learning process.

Orcid ID: https://orcid.org/0000-0002-1314-2441

To cite this article: Jorge Díez (2021): Lack of interest? Self and peer assessment as a means to improve students' engagement, Innovations in Education and Teaching International, DOI: 10.1080/14703297.2021.2013288

To link to this article: https://doi.org/10.1080/14703297.2021.2013288

Lack of Interest? Then Correct Exams! Methodology for Students' Engagement

In their first year of university, students perceive some courses as least related to their degree, and hence, show a lack of motivation in them, dedicating more time to those that they consider as more related to their degree. This leads to a decrease in students' performance in these courses, even when they have sufficient capacity to do well. Moreover, it is observed that students are not aware of their knowledge level regarding the course content and generally overestimate it. This paper presents a method to increase students' engagement in parts of the courses that are most difficult. This method was tested in one course, and the results support its effectiveness.

Keywords: self-assessment; peer assessment; feedback; engagement; first year

Introduction

Several theories have focused on motivation, which is an important factor for the successful learning of a new subject (Seifert, 2004). For university students, finding motivation is easy as they generally pursue degrees that interest them. However, in their first year at university, students must take some mandatory courses that they consider as outside their interests because they are not directly related to the field of their degree.

The lack of motivation of students for these courses, which are included more in the first year syllabus, shows that many do not pay sufficient attention to them throughout the semester (focusing on other courses that are more related to their field of interest). Getting students more involved in these courses is an arduous task with no simple solution.

This paper will present a way of using self-assessment and peer assessment to make students more involved in subjects that they are less familiar with. To this end, these techniques will be applied in a course of these characteristics, "*Fundamentals of*

Computer Science," which carries six ECTS (European Credit Transfer System) credits in the first semester and year, and it is common to all the degree programs taught at the Polytechnic School of Engineering of Gijón. This school offers degrees in computer science, telecommunications, and industrial engineering degree (eight degree programs), with common classes and examinations for all degrees in that course. This course is closely related to the computer science degree, and that is how students perceive it. However, students from other degrees perceive it as unrelated to their field of interest and do not pay the necessary attention to it.

The contents have been designed to cover the specific competence: "basic knowledge on the use and programming of computers, operating systems, databases, and software with application in engineering." A major part of the teaching activities are devoted to teaching the basics of programming, and many are not motivated as they are not sure of its practical relevance. Moreover, learning programming is not easy (Jenkins, 2002). Furthermore, initially, many industrial engineering students see computers as a tool for designing and preparing documents and are not interested in acquiring that knowledge. Hence, they do not realize the various advantages of learning to program.

TABLE 1: CLOSE TO HERE

The author has taught this course for several academic years, from 2011 to 2016 as part of the computer science degree and from 2016 to 2018 as part of the industrial engineering degree (in a theory group). Table 1 shows the aggregate statistics of the continuous assessment for the courses in the January term, which is continuously assessed. Hence, in the five academic years taught in the computer science degree in the January term, 79.50% were presented and 51.31% (65.02% of those presented) of enrolled students passed it. In the following two academic years, the same teacher taught this course (with the same contents and evaluation mechanism) in the industrial engineering degrees and 88.69% of the students took the January exam and 44.06% passed (39.08% of those enrolled). The performance of students in industrial engineering degrees was worse than that of those in the computer science degree due to a lack of motivation in the former. We think in that way because the course is a first year and semester course and does not require knowledge beyond that acquired in the high school.

Notably, the percentage of students taking the January exam is higher in the industrial engineering degrees than in the computer science degree, suggesting that students are not aware of the knowledge level of the course and they probably overestimate their capabilities when taking the final theory test.

To make the students of other programs, who are unfamiliar with computer science, become more involved in the course and more aware of their knowledge level, a teaching methodology was designed to enable them to anonymously compare their knowledge with their classmates and motivate them in the most complicated part of the course: programming. This is the most extensive part of the course in which students tend to have the most problems.

This methodology, explained in the Materials and Methods section, is based on the use of two different assessment mechanisms: self-assessment and peer assessment. Some instructors believe that applying these methods in the first academic year is too early because students may not be mature enough; however, as can be seen from the literature review by Nulty (2011), the use of self-assessment and peer assessment is highly recommended in the first year of university studies. Orsmond et al. (2004) have used both mechanisms in the same course to assess student-made posters, showing that they help students achieve their learning objectives. In this paper, we show how both mechanisms are used to evaluate a midterm exam and how it influences student learning. There is extensive literature listing the benefits of this type of correction in which students participate (Falchikov & Goldfinch, 2000; Gillanders et al., 2020; Panadero & Brown, 2017; Wanner & Palmer, 2018).

Regarding self-assessment and peer assessment, the aim is not to change the assessment methodology of the course (students will continue to be assessed by the teacher) but to involve students in the evaluation of a midterm exam by assessing themselves and some peers. Hence, students will dedicate more effort to learning the subject when it is crucial for the correct follow-up of the course (during the programming topic). This method leads to greater involvement of the students, whose motivation and knowledge of the course content will be high.

Material and methods

The proposed methodology will be applied to students of the *Fundamentals of Computer Science* course of industrial engineering degrees. The activity was conducted with students in one of the 12 theory groups. This group comprised 54 students who were offered the opportunity to participate by showing them the improvement in understanding of the course contents and the possibility of anonymously comparing themselves with their peers. Participation was voluntary and without incentives, as it was unclear how incentives could be included without altering the normal progression of the course (Double, 2020). Finally, 37 students participated in all requested corrections. At all times, students were made aware of the importance of this experience, to motivate them to make good corrections because the experience would lose its relevance if the corrections were made with a lack of interest (Turner, 2011).

The course has four parts, with the "*Introduction to programming*" having the most weightage, as half of the course content is programming. There are two exams on programming during the continuous assessment, one when half of the syllabus is covered and the other one at the end of the course. The exams are graded on a scale of 0 to 10, and a score of five or more indicates that the student passed the exam. The first exam of the programming part is a critical point for the course follow-up. Our aim is to act at that moment, so that those students who have obtained poor results are aware of their mistakes, know how to solve them, better understand the course content, and can continue with greater guarantees of passing it.

Therefore, the exam was conducted similarly to the previous years: the students had to create three small programs on a computer using the programming tools taught in the course. To achieve the objectives, a procedure was devised to evaluate each student's midterm exam using different actors: i) the student (twice: once at the beginning and once at the end of the procedure), ii) several classmates (the student himself acting as an evaluator of the work of other classmates), and iii) the teacher.

After taking the exams, students submit their solutions by sending text files containing the programs. However, for evaluation the submitted solutions are provided as image files to the students. This is because it is intended that during the assessment, students will read the source codes and try to understand them, instead of simply running the programs to see if they work.

After the exams, students are provided with a rubric showing a possible solution to the problems posed and a set of guidelines on how the exams should be assessed. These guidelines state that points are lost for mistakes since, according to Gillanders et al. (2020), *loss aversion* is more effective than indicating points for success. Subsequently, they were asked to evaluate their exams according to these guidelines (Phase 1 of Figure 1).

FIGURE 1: CLOSE TO HERE

After this first phase of self-evaluation, each student was given three completely anonymized classmate exams and asked to assess them using the rubric (Phase 2 of Figure 1). At the end of this second phase, students will only be aware of the grade they gave to themselves and the grades they gave to the other students.

With the experience gained from the corrections in the first and second phases, students were asked to re-evaluate their exam (Phase 3 of Figure 1). In the third phase, it is possible that students change their original grades after becoming aware of the performance of other students and gaining some experience in correction. Having seen other possible solutions, detected possible bugs and bad programming practices in the other exams, they will have a better knowledge of programming and therefore can have a more critical view.

Finally, an anonymized list will be published with the scores: the one granted by the teacher and the three scores given in the peer assessment process. The self-assessment grades are not made public, and only the individual knows it.

Results and discussion

Table 2 shows the performance obtained in the January call after applying the proposed methodology. The percentage of students who presented in that call was slightly reduced, and the percentage of those who passed compared to those enrolled increased significantly. Therefore, the presented methodology had the desired impact way, and those studying industrial engineering degrees passed the course in similar percentages as those of the computer science degree. The slight decrease in the percentage of presented students could be because the methodology applied has made the students see their knowledge level regarding the course content.

TABLE 2: CLOSE TO HERE

FIGURE 2: CLOSE TO HERE

Figure 2 shows the variation in scores between the two self-assessments against the average score obtained in the peer assessment. For example, a student who was graded 1.1 in the peer assessment had reduced his score by over three points in his second self-assessment. Not all students showed such marked differences. Among them, 70.27% (26 out of 37) had lowered their grade in the second assessment, 10.81% (4 students) did not alter their grades and the remaining 18.92% (7 students) felt they deserved a better rating.

The results show that students analyzed their work more critically in the second self-assessment after assessing others' work. Examining other exams in detail led students to find mistakes that they had not previously detected in their work (Hanrahan & Isaacs, 2001). A more critical view of their knowledge level regarding the course content may have influenced the decrease in the percentage of students presented in January, as shown in Table 2. Having seen the exams of their peers and having had to grade them led to a posteriori study of the course content, helping them be more critical of their exams (Hanrahan & Isaacs, 2001; O'Donovan et al., 2004).

FIGURE 3: CLOSE TO HERE

Since there are four scores for each exam (the teacher's score, two self-assessments, and the average obtained in the peer assessment), these scores were compared in aggregate terms. Figure 3 shows the distribution of the grades obtained by the exams in the four corrections. The students, when they evaluated their exam for the first time, were graded mostly with grades from 5. However, in the second self-assessment, the number of failures (scores below 5) increased. It is striking that neither the first nor the second self-assessment had a single student score in the range [0, 2]. Interestingly, in line with studies such as Falchikov & Goldfinch (2000), the distribution of grades obtained by peer assessment is similar to that obtained by the teacher's assessment.

TABLE 3: CLOSE TO HERE

To analyze the extent of similarity between the grades given by the teacher and those obtained with the intervention of the students, the correlations and the mean absolute error (MAE) were calculated. Table 3 lists the results. It shows that the correlation between teacher and student ratings in the first self-assessment was the lowest. It increases significantly in the second, which corroborates the above analysis that the students checked their work more critically the second time. The second self-assessment and peer assessment showed similar correlations. For the MAE, the students initially gave themselves scores that were on average 2.05 points away from the grade given by the teacher. In the second self-assessment, the mean of the differences in the rating was reduced to 1.56, showing that students were more aware of their knowledge of the course content. Regarding peer assessment, the MAE was 1.37, implying that the score obtained by peer assessment was, on average, 1.37 points higher or lower than that given by the teacher. Therefore, although the teacher and peer assessment grades had similar distributions, the MAE shows that the peer assessment grades in this experiment are not accurate. The grades might improve if students assess more exams to make the mean more meaningful; however, this is not a practical solution as students would be burdened with more work.

Conclusion

It can be concluded that the proposed methodology has improved the performance of industrial engineering students. The improvement in the results is not comparable to the result obtained in the computer science degree in these five years.

Nevertheless, there is a marked improvement compared to the results of the two previously analyzed academic years in the industrial engineering degrees.

Additionally, students are more critical of their knowledge level regarding the course content, which is reflected in the results of the second self-assessment as the majority of students reduced their grades. The percentage of students presented in the January term decreased while the percentage of passing grades increased, indicating that the number of students who stopped presenting owing to the awareness about their lack of preparation has increased.

Finally, the proposed methodology could be incorporated into other courses that face similar problems. However, over-dependence on this methodology could increase the workload for students and teachers (Hanrahan & Isaacs, 2001).

Therefore, this methodology is effective, and it is recommended to use it promptly in parts that are difficult for the students.

Declaration of interest statement

The author does not report any potential conflict of interest.

References

Double, K.S., Mcgrane, J.A., & Hopfenbeck, T.N. (2020). The impact of peer assessment on academic performance: A meta-analysis of control group studies. *Educational Psychology Review*, *32*(2), 481-509.

Falchikov, N. & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, *70*(3), 287-322.

Gillanders, R., Karazi, S., & O'Riordan, F. (2020). Loss aversion as a motivator for engagement with peer assessment. *Innovations in Education & Teaching International, 57*(4), 424-433.

Hanrahan, S.J. & Isaacs, G. (2001). Assessing self-and peer-assessment: The students' views. *Higher Education Research & Development*, *20*(1), 53-70.

Jenkins, T. (2002). On the difficulty of learning to program. In Proceedings of the 3rd annual conference of the LTSN Centre for Information and Computer Sciences *Vol. 4*, 53-58.

Nulty, D.D. (2011). Peer and self-assessment in the first year of university. *Assessment & Evaluation in Higher Education, 36*(5), 493-507.

O'Donovan, B., Price, M., & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education*, *9*(3), 325-335.

Orsmond, P., Merry, S., & Callaghan, A. (2004). Implementation of a formative assessment model incorporating peer and self-assessment. *Innovations in Education & Teaching International, 41*(3), 273-290.

Panadero, E. & Brown, G.T.L. (2017). Teachers' reasons for using peer assessment: Positive experience predicts use. *European Journal of Psychology of Education*, *32*(1), 133-156.

Seifert, T. (2004). Understanding student motivation. *Educational Research, 46*(2), 137-149.

Turner, S.A., Pérez-Quiñones, M.A., Edwards, S.H., & Chase, J. (2011). Student attitudes and motivation for peer review in CS2. In Proceedings of the 42nd ACM technical symposium on Computer science education. 347-352.

Wanner, T. & Palmer, E. (2018). Formative self-and peer assessment for improved student learning: The crucial factors of design, teacher participation and feedback. *Assessment & Evaluation in Higher Education*, *43*(7), 1032-1047.

Table 1. Aggregate statistics of results in computer science and industrial engineering degrees.

Degree	Academic courses	Number of students	Number of students per course	Students presented in the January call	Passed out of presented	Passed out of enrolled
Computer Science	5	310	62	79.50%	65.02%	51.31%
Industrial Engineering	2	133	66.5	88.69%	44.06%	39.08%

Table 2. Student performance after innovation implementation.

Degree	Academic courses	Number of students	Number of students per course	Students presented in the January call	Passed out of presented	Passed out of enrolled
Computer Science	5	310	62	79.50%	65.02%	51.31%
Industrial Engineering	2	133	66.5	88.69%	44.06%	39.08%
Industrial Engineering after experience	**1**	**54**	**54**	**85.18%**	**60.87%**	**51.85%**

Table 3. Correlation and mean absolute error (MAE) between teacher-provided grades and grades obtained by student participation through self and peer assessment.

Degree	**Correlation**	**MAE**
First self-assessment	0.6450	2.05
Second self-assessment	0.8175	1.56
Average peer assessment	0.7887	1.37

Figure 1. Methodology applied. In Phase 1 students self-assess their exams, in Phase 2 they grade the exams of three classmates and in Phase 3 they self-assess again. In parallel, the teacher corrects the exams of all students.

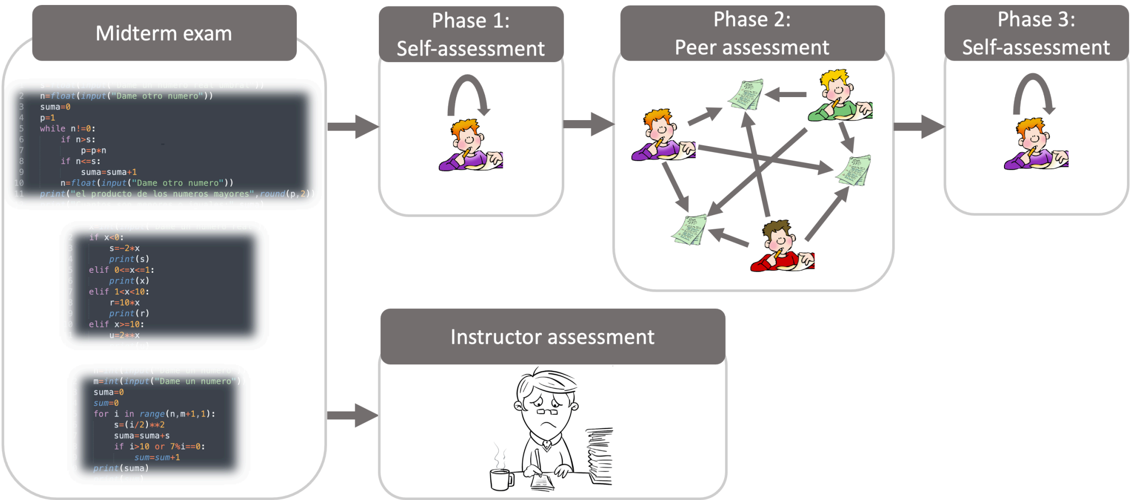

Figure 2. Difference between the grades obtained in the two self-assessments versus the average grade obtained in the peer assessment. The exams are graded on a scale from 0 to 10 and negative values indicate the decrease in the grade in the second self-assessment. The trend line and the equation of the line are also shown together with the R^2 value.

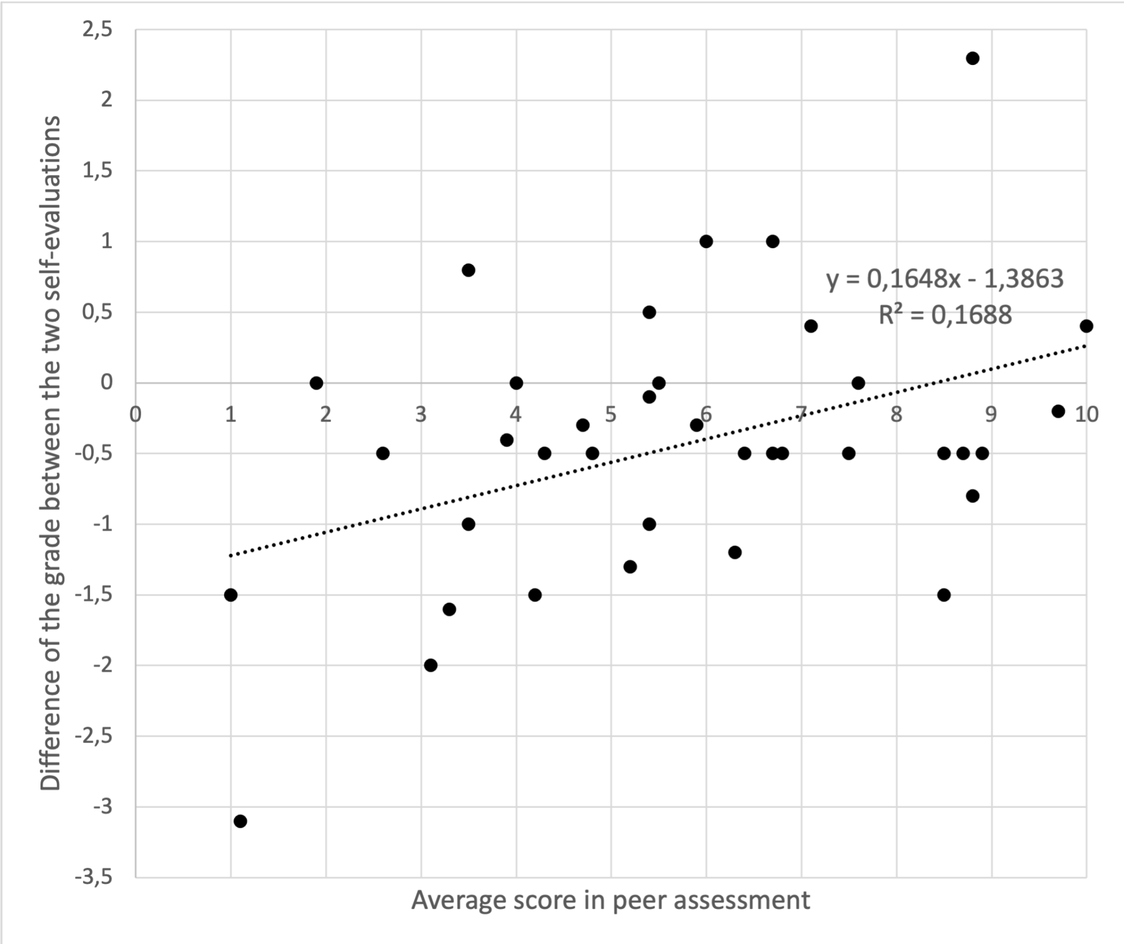

Figure 3. Distribution of the examinations according to the grade obtained in each of the assessments.

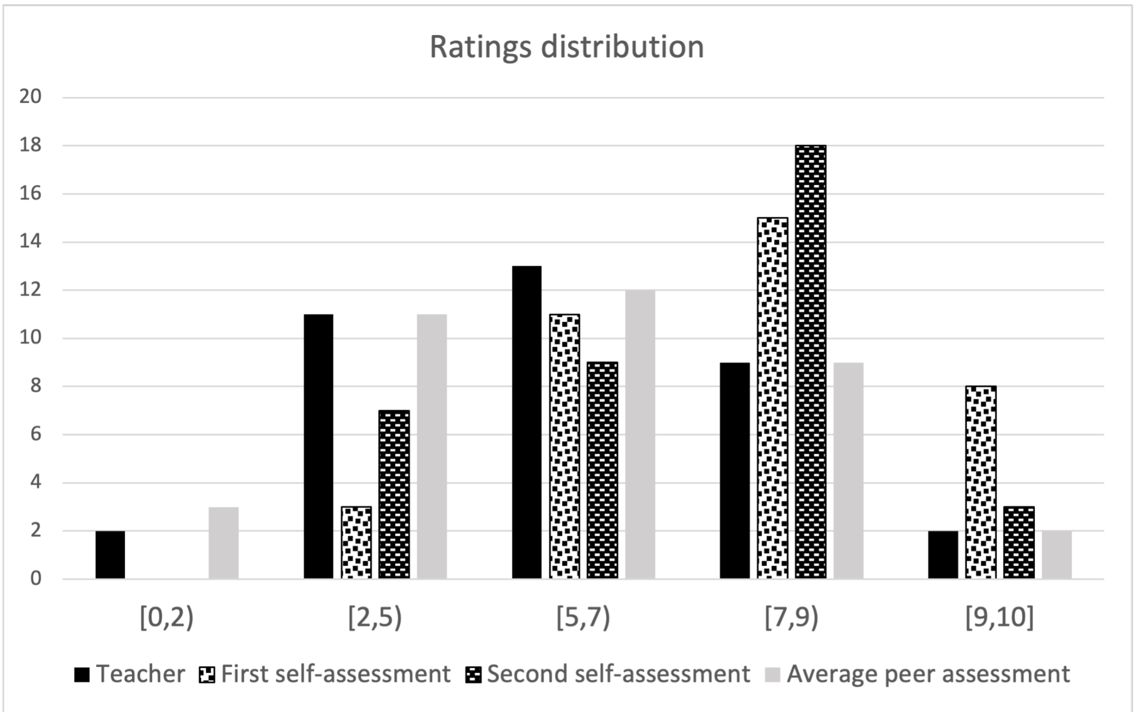

Figure 1. Methodology applied. In Phase 1 students self-assess their exams, in Phase 2 they grade the exams of three classmates and in Phase 3 they self-assess again. In parallel, the teacher corrects the exams of all students.

Figure 2. Difference between the grades obtained in the two self-assessments versus the average grade obtained in the peer assessment. The exams are graded on a scale from 0 to 10 and negative values indicate the decrease in the grade in the second self-assessment. The trend line and the equation of the line are also shown together with the R^2 value.

Figure 3. Distribution of the examinations according to the grade obtained in each of the assessments.