Online peer assessment to improve students' learning outcomes and soft skills

Valutazione tra pari online per migliorare i risultati dell'apprendimento e le soft skill degli studenti

Daniela Amendola* and Cristina Miceli

School of Biosciences and Veterinary Medicine, University of Camerino, Italy, daniela.amendola@unicam.it*; cristina.miceli@unicam.it

*corresponding author

ABSTRACT A central objective in the university context is to promote innovative teaching activities that increasingly see students as an active part of the learning process, fostering their learning outcomes, knowledge, and soft skills. The use of online technologies has the potential to support and innovate frontal teaching lessons, also solving problems associated with lack of time and space in the standard lesson environment. To achieve the above objective, we included the collaborative activity of peer assessment within the online Genomics laboratory as a formative process for the second consecutive year. This experimental procedure was conducted entirely online, using the Moodle e-learning platform. The results reported in this paper are based on qualitative and quantitative analysis of the peer assessment process and on a questionnaire on the students' subjective perception of this collaborative activity.

KEYWORDS Peer Assessment; Formative Evaluation; Moodle; E-Learning; Moodle Workshop Module.

SOMMARIO Nel contesto universitario il principale obiettivo è garantire una didattica innovativa che veda sempre più lo studente come parte attiva nel processo formativo, favorendo il miglioramento degli apprendimenti in termini di conoscenze e di soft skills. L'utilizzo delle tecnologie online permette di supportare ed innovare le lezioni frontali, anche risolvendo i problemi di mancanza di tempo e spazio connessi con la didattica convenzionale. Per raggiungere i suddetti obiettivi, per il secondo anno consecutivo, grazie all'utilizzo delle tecnologie digitali, abbiamo inserito all'interno del percorso laboratoriale online di Genomica l'attività collaborativa di peer assessment, come processo valutativo con funzione formativa. Questa procedura sperimentale è stata interamente condotta online, utilizzando la piattaforma e-learning Moodle. I risultati del presente lavoro si basano sull'analisi qualitativa e quantitativa del processo di peer assessment e di un questionario sulla percezione soggettiva degli studenti relativa a tale attività collaborativa.

PAROLE CHIAVE Valutazione tra Pari; Valutazione Formativa; Moodle; E-Learning; Modulo Workshop di Moodle.

1. INTRODUCTION

Peer assessment (also called peer review) is a collaborative learning technique based on a critical analysis by learners of a task or artefact previously undertaken by peers (Persico, Pozzi, & Sarti, 2009). In the peer assessment process, students reciprocally express a critical judgment about the way their peers performed a task assigned by the teacher and, in some cases, they give a grade to it. According to the literature on collaborative learning, the peer assessment process supports and enhances learning in both the students who receive the feedback and in those who give it, because the activity triggers self-assessment and critical reasoning with a focus on the artefacts produced by both (Kreber, Anderson, Entwhistle, & McArthur, 2014). To produce the feedback, the students can use a rubric (Hafner & Hafner, 2010), which is a schema of the criteria for assigning marks for each step of the task. The rubric is usually prepared by the teacher in collaboration with the students themselves, thus promoting metacognitive reflection on the quality of the task or artefact to be produced. Generally, the peer assessment process entails that the students provide their peers with detailed qualitative feedback to guide and help them in the constructive revision of their work.

Previous research on peer assessment (Nicol, Thomson, & Breslin, 2013) has shown that producing reviews with feedback engages students in generating evaluative judgement, both about the work of peers and about their own work. The capacity to produce quality feedback is a fundamental skill that should be developed in graduate students, because it is an essential component of their self-regulated learning abilities. In particular, through peer assessment, students enhance their self-assessment skills and their understanding of the main concepts treated during the collaborative activities (Chao, Saj, & Hamilton, 2010). In addition, students reported developing a deeper understanding of the course material thanks to analysis of the perspectives of their peers (Guilford, 2001). Usually, peer assessment is used with written assignments, wherein students provide feedback which may be formative only or both formative and summative.

However, outcomes from the peer assessment process depend on student behaviours and understanding. Outcomes in terms of student perception were found to be positive in some studies (Chao et al., 2010; Nicol et al., 2013; Landry, Jacobs, & Newton, 2015; Amendola & Miceli 2016) but less positive in others. Among the former, Venables and Summit (2003) reported that the majority of the students felt they had greatly improved learning of the topics proposed, considering the feedback from their peers, with 26% of the students changing their opinion from negative to positive on the effectiveness of the peer evaluation activity at the end of the process. Among the latter, Mulder, Pierce and Baik (2014) showed a negative change in the perception of the students toward the peer assessment after their experience in an interdisciplinary investigation. Bangert-Drowns, Kulik, Kulik and Morgan (1991) observed that scepticism about the ability of peers to give accurate feedback tends to inhibit the learning process.

More recently, the impact of innovative assessment on student experience in higher education was reported (Bevitt, 2015). Interestingly, this article reports that innovation is a concept defined individually by students themselves, shaped by diverse assessment experiences and preferences, and therefore the resulting impact is difficult to predict. Aquario and Grion (2017) investigated peer evaluation processes in the university system, with a focus on student perception of evaluation; they reported very positive perception by more than 80% of the students.

In summary, the beliefs and perceptions of students about the peer assessment process still need further investigation at different levels and on different aspects. Related investigations are currently being carried out by other research groups, as reported in Grion, Serbati, Tino and Nicol (2017).

Another important aspect of the peer assessment process is whether the grades and judgements given by peers are consistent with each other and with those of the teacher. This analysis can give indications on students' level of understanding of the topics being studied and on their capability to provide useful help with

their feedback. Topping, Smith, Swanson and Elliot (2000) observed that peer assessment at postgraduate level did not show significant deviation with respect to teacher evaluation. The reliability and accuracy of students as assessors of their peers, however, seems to be higher at postgraduate level than undergraduate level (Hongli, Yao, Xiaojiao, Mindy, Youngsun, Kyung, & Hoi, 2015). In this respect, the contribution made in the present paper regards peer assessment activities undertaken in scientific subjects during a Master's course.

The research questions we address are:

- To what extent is peer evaluation accurate and reliable in comparison with the grades given by the teacher?
- Is there an improvement in the artefacts produced by the students after the peer assessment process?
- What is the students' behaviour and perception with respect to the peer assessment activity?

2. CONTEXT OF THE STUDY

In this paper, we describe and analyse an experience of online peer assessment involving students attending the first year of the Master's degree course in Biological Sciences at the University of Camerino (academic year 2016/2017). Specifically, this regarded a collaborative activity called "Online Practical Simulation Exercises in Genomics" (in the following: Genomics laboratory) based on a online lab simulation. In this context, peer assessment was deemed a suitable collaborative technique because of its potential for fostering deeper understanding of the course contents, for improving metacognitive skills, and for evaluating and promoting collaborative skills and critical thinking (Stevens & Levi, 2005). The online simulation was delivered through the Moodle e-learning platform and supplemented the theoretical lessons in class. The purpose was to offer an integrated learning environment bridging the theoretical lessons and the practical collaborative activities in order to foster students' learning outcomes and soft skills. Examples of soft skills are critical thinking, responsibility and team building, which are useful for students' learning and job placement.

This path allowed the students to perform analyses of genetic sequences based on real experimental data; the sequences were not yet available in databases because they belong to an ongoing project of annotation of new genomes from eukaryotic microbes.

On the Moodle e-learning platform, students could access informative training material and four video tutorials. The video tutorials were prepared by the teacher to provide a guide for the use of online software, databases, and open source software for gene sequence analysis. The tutorials presented a case study of one gene-sequence as an example. After a period of individual study, we asked the students to solve a problem consisting in the analysis of a new gene sequence (not available in the databases), including identification of the potentially encoded protein and alignment of the new protein with evolutionary related proteins in the databases.

The online project comprised an initial phase in which each student worked alone and a second part in which they engaged in collaborative activities assigned by the teacher that entailed a peer assessment process using the Moodle Workshop tool.

To test the efficacy of our online laboratory with collaborative peer assessment activities, we analysed the qualitative and quantitative data extracted from the Moodle platform, focusing on the peer assessment process. The approach adopted was a formative assessment one: the students formulated their peer assessment with the guidance of a rubric containing a list of evaluation criteria and the authors of the artefacts were given the possibility to revise their work using the suggestions received before submitting it for the final evaluation by

the teacher. More specifically, a rubric consists of a set of clear expectations or criteria used to help teachers and students focus on what is valued in a subject, topic, or activity (Airasian & Russell, 2008). In our case, the rubric was discussed and prepared by the teachers in conjunction with the students during the preparatory phase of the peer assessment, which comprised two three-hour face-to-face meetings. The students also provided qualitative, critical and constructive feedback to help their peers improve their work and corresponding learning outcomes. The same rubric was used by the teacher for the final summative evaluation. To measure the learning improvements generated by the peer assessment, the teacher also performed an evaluation of students' work before the peer assessment process, adopting the same evaluation criteria; this produced information and data on the starting level of the students when working alone. The decision to allow revision after the peer assessment was suggested by the literature: "The performance seems to improve further when the evaluation processes include feedback and the opportunity of revision" (Gibbs & Simpson, 2004).

3. METHODOLOGY

In this section, we describe in more detail the method used for the Genomics laboratory, as well as the data collection methodology adopted. The flow of the teaching and research activity is shown in Figure 1.
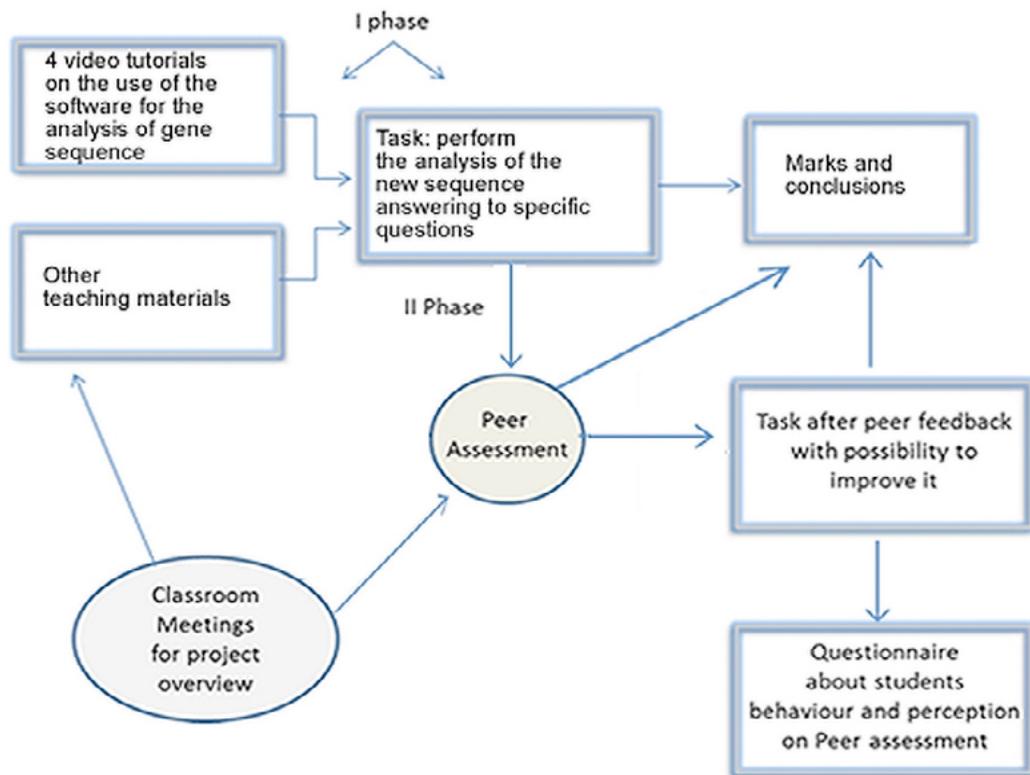

Figure 1. Flow of the teaching and research activity.

3.1. *Participants*

The study was conducted with a total of 45 Italian and International students (26 females and 19 males)

divided into 8 groups, with 5 or 6 students in each. Participation was voluntary and the students actively participated in the online activities.

3.1.1. *Online learning activity*

The Genomics laboratory was supplementary to classroom activities. It allowed students to analyse genetic sequences from real data deriving from an ongoing project of annotation of new genomes from eukaryotic microbes, specifically, ciliated protozoa of the genus Euplotes. The students accessed to training materials and video tutorials offering guidance in the use of online software, databases and other open resources for sequence analysis. The tutorials included a gene-sequence case study as an example.

At the outset of the activity, the students were divided into groups of five or six students. For each group, an appropriate sequence was selected by the teacher. After consulting the tutorials, the students were instructed to work individually on analysis of the sequence assigned to their specific group. The activity was carried out in four steps, guided by four prompts. Each step involved the use of software or specialised procedures. Once the task was concluded, each student was asked to upload their report using the Moodle Workshop module. Then the second phase commenced, entailing interaction with the other students in the same group.

3.1.2. *Collaborative online activities*

To reproduce the collaborative environment of the school laboratory, students worked in groups during the online collaborative activities and mutually assisted each other thanks to their different levels of competence and knowledge. This is when soft skills, such as critical thinking, responsibility, and time management can be developed.

To this end, the "Workshop" module of the Moodle platform (Figure 2) allowed the collection, review and peer review of the work done by students. This module allowed teachers to organize, carry out and monitor the various stages of the peer assessment activity: presentation of the procedure path, anonymous distribution of the work done by the students for evaluation, distribution of the rubric for feedback and peer assessment, and tracking of the activities.

During the peer assessment activity each student uploaded the first draft of their report using the Workshop tool, then provided anonymous feedback to peers (in our case, two other students) by filling in the rubric with quantitative and qualitative judgement of the work performed (Figure 3). In this way, each student received feedback from two peers.

For the quantitative judgement, we asked students to assign an overall score as a percentage. For the qualitative evaluation, we asked them to give feedback with tips that could help peers in improving their tasks, before delivery for the final evaluation by the teacher.

Figure. 2. Screenshot of the Workshop module on the Moodle platform

The students had 47 days to complete all the activities. This period comprised: 2 days (3 hours per day) for the introductory meetings, 20 days for individual study of the materials and for producing the first draft of the task, 15 days for the peers' task evaluation, and 10 days for review and resubmission of the final version of their outputs.

```
Rubric grid for feedback
Answer to the following questions:

Questions with marks(give a mark from 1 to 10 where 10 is the maximum score for each question)
1.Is the analysis of the open reading frames well described ?..../10
2.Is it clear how the intron is predicted?..../10
3.Is the experimental plan to validate the intron well described? ...../10
4.Does the author incorporate comments to the aligned proteins in a stimulating way? ..../10
5.Are there any grammar, spelling punctuation, etc. mistakes? ..../10
6.Is the report well organized and does it follow a logical train of thought? ..../10
7.Is the writing concise and are the sentences directed to the point? /....10
8.Is the language appropriate? ..../10
9.Does the author use appropriate scientific and academic terminology?.... /10
10.Does the author answer to all the questions in an adequate manner?.../10
```

Open question
Give your qualitative feedback writing what is well done, what could be improved and was done incorrectly:..

Figure 3. Evaluation rubric used by the students and teacher in the online Genomics computer laboratory.

To resubmit their final versions, the students used a different module of the Moodle platform called "Task".

3.2. *Data collection methodology*

The results of the study are based on the data extracted from the Moodle platform, the analysis of the work carried out by the students, and the analysis of the results of a qualitative/quantitative questionnaire on the students' subjective perception of the learning activity.

Specifically, we analysed: *(i)* the peer evaluations to verify their accuracy and reliability (comparison between the evaluation of the two peers and comparison between the average of those evaluation and the evaluation of the teacher on the first version of the task); (*ii*) the teacher's evaluation of the first version and final versions, to verify any improvement in the students' artefacts after the peer assessment process; (*iii*) the data collected through a qualitative/quantitative online questionnaire on the students' experience and perception about the peer assessment activity, proposed to the students at the end of the laboratory activity. As for (i) and (ii), to verify the possible correlation between our variables we performed a linear regression analysis, utilizing a fit function as $y(x)=a*x+b$, to derive the Pearson correlation with bootstrap analysis to determinate the degree of accuracy of the model used with our sample. For our analysis, we considered the following variables: average grade between the grades of peer 1 and peer 2 versus the grade of the teacher; the grade given by the peer 1 versus grade given by peer 2. Our aim was to establish which functional correlation is present among the variables, as described above. We analyzed the available data with the help of graphical representations and statistical processing, using the statistical functions of Microsoft Excel. The

data analyzed came from a sample of 45 students.

As for (iii), the questionnaire consisted of 21 questions classified using the 5-point Likert scale (Likert, 1932) and 10 open-ended questions. All the 45 students responded to the questionnaire. The answers to the open-ended opinion questions were analysed for common themes, and major advantages and disadvantages of the experience were categorized.

4. RESULTS AND DISCUSSION

To answer the first research question "To what extent is the peer evaluation accurate and reliable in comparison with the teacher's grades?" we started with analysis of the results on the comparison between the grades given by the teacher and the average of the grades given by the peers. In Figure 4, the x-axis shows, for each task done by a student, the grade assigned by the teacher, while the y-axis corresponds to the arithmetic average of the grades given by the two peers. We have 45 tasks and in Figure 4 each point corresponds to a task carried out by a student and evaluated by the peers and the teacher.

In order to obtain statistical information on the collected data, in Figure 4 we report the linear regression of the data (blue line), using $y(x)=a*x+b$ as a fitting function. The resulting fitting parameters are a=0.75 and b=19.1. From the regression line and its distance from the data, we obtained the correlation coefficient, equal to r=0.82. Being r=0.82 > 0.7, the linear correlation between the data appears strong. Therefore, we can conclude that the two peer assessments correspond well with the grades that the teacher assigned. Indeed, all the data are within a range of +/-20% variation with respect to the grade of the teacher. A possible improvement in the assessment procedure based on peer assessment could be to request a third peer assessment if the grades of the first two peers diverge. This is discussed below.
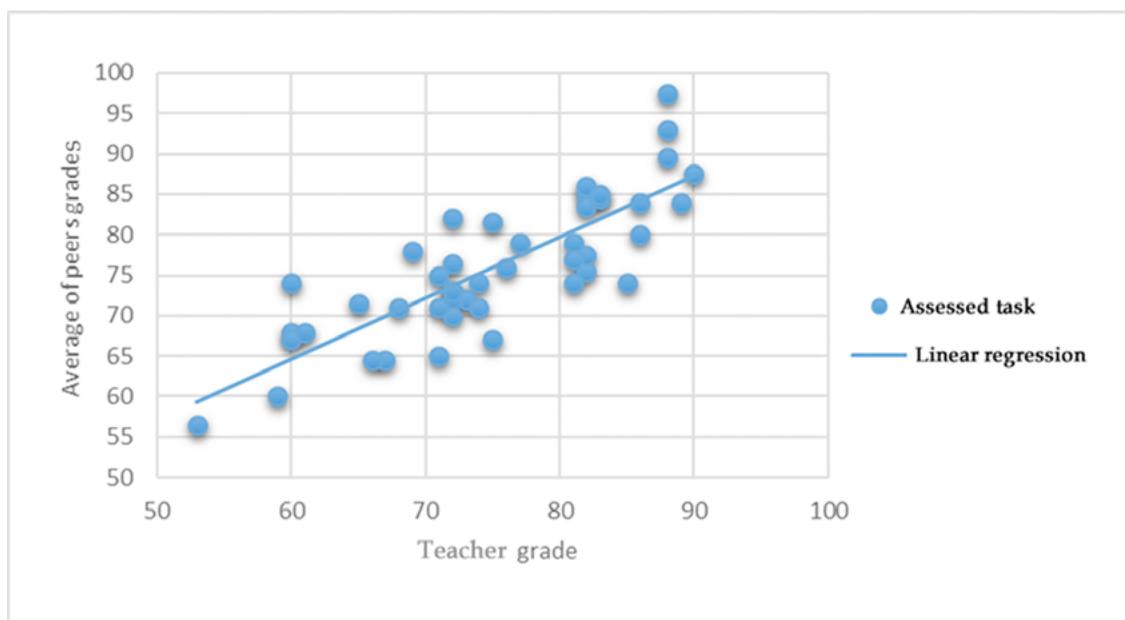

Figure 4. Regression analysis of the average of two students' reviewer grades vs teacher grades.

In Figure 5 we report the grades assigned by peer 1 on the x-axis and by peer 2 on the y-axis. Each point in

Figure 5 corresponds to a single assessed task, with an overall score in hundredths and 45 double assessments available.

From a preliminary analysis it is evident that there is a greater dispersion of data compared with Figure 4, even if we observe that increased peer 1 grades are matched by increased peer 2 grades.

To examine the data in depth, we used $y(x)=a*x+b$ as a fitting function, where the resulting fitting parameters are a=0.70 and b=23. In this case, the distance between the data and the regression line (orange) generated a Pearsons' correlation coefficient r=0.57, which represents a moderate correlation between the variables (r being in the range of $0.3 < r < 0.7$).

The result of a moderate correlation between the variables suggests that in the cases in which there is an excessive deviation between the grades of the peers (for example greater than 30%), assessment from a third peer would be required, exactly as happens with the peer review process in scientific fields. Indeed, as we will see from the results of the perception questionnaire reported below, the majority of the students considered it would be useful to receive more than two reviews from peers, in agreement with the position mentioned earlier.
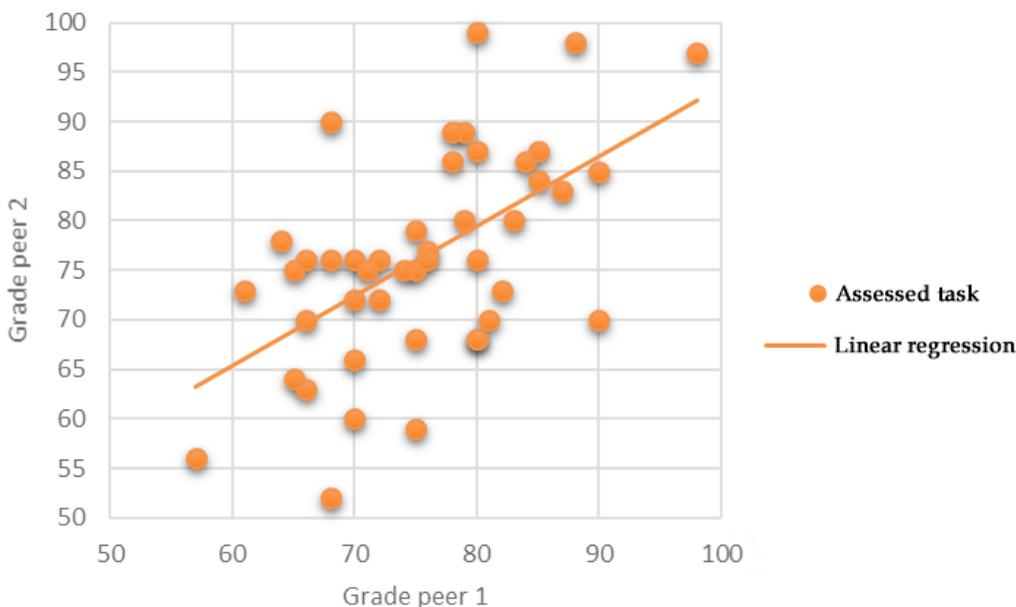

Figure 5. Regression analysis of the two student reviewer grades.

In answer to the second research question "Is there an improvement in the artefacts of the students after the peer assessment process?" in Figure 6 we show the variation of the grades awarded by the teacher before and after the peer assessment process. The clear and significant quantitative result is that in 38 cases over a total of 45 the grade that the teacher assigned to the first draft increased with a maximum 8 point improvement.

In only seven cases the grade remained unchanged (2, 3, 4, 9, 29, 32, 41 in Figure 6).

Furthermore, we found that, with few exceptions, the highest grade variations regard the student products that had already received a high grade in the first assessment done by the teacher. Regarding products for which the first draft version grade is lower than or equal than 60, we found a significant improvement in only one case. Therefore, students who were very weak at the beginning of the course activity showed little improvement. Different results for improvement by low-performing students were found by Lan and Fei

(2015). They found that the low- and average-achieving students performed significantly better immediately after the integration of a peer assessment activity. Moreover, their model seemed to have had less impact on the performance of high-achieving students. Similar conclusions were also reached by Li (2011) in a previous study. As mentioned in discussion of Figure 6, we agree on the significant impact peer assessment has on average-achieving students, but by contrast we found less impact on low- and very low-achieving students and a positive significant impact on high-achieving students.
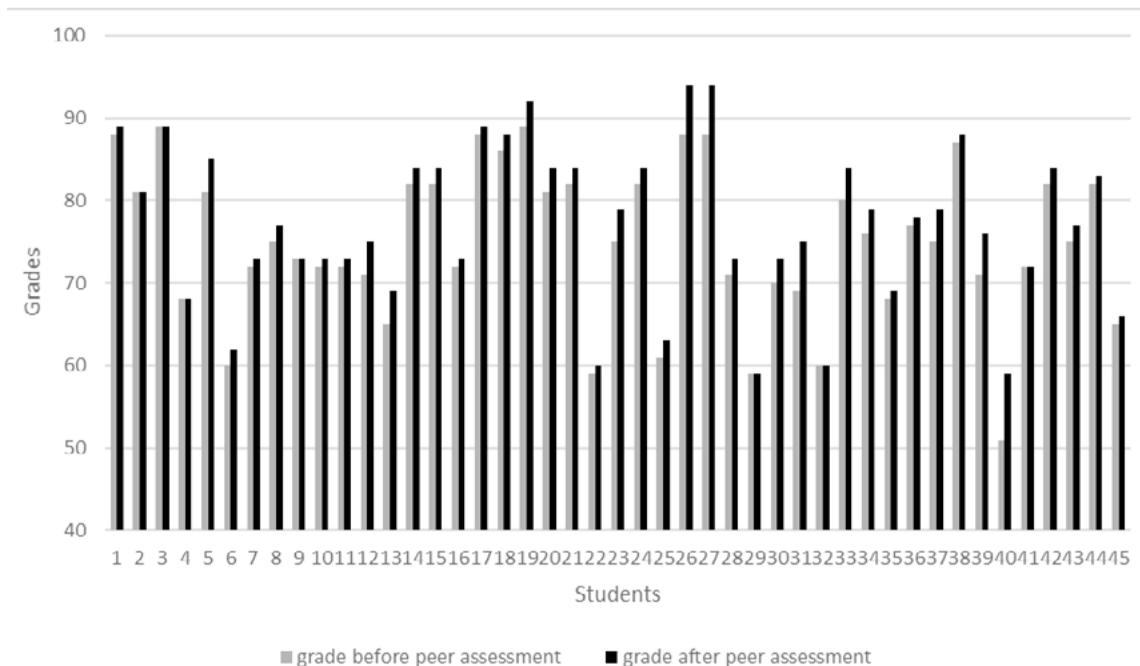

Figure 6. Comparison between teacher-assigned grades for the draft version (light columns) and the final version (dark columns) of the assignment.

We now move on to answer the third research question "What is the students' beliefs and perception about the peer assessment activity?". In the following we report a few examples of the questionnaire results (classified using a 5-point Likert scale) proposed to the students about their behavior and perception on peer assessment. We also evaluated the mean value (MV) of the responses associating a scale from 1 to 5 to the five possible choices given in the questionnaire (SD=1; D=2; N=3; A=4; SA=5) (Table 1). Forty-five students answered the questions: 58% agreed (A) or strongly agreed (SA) that they had used their peers' feedback to revise their tasks (MV=3,45); 68% felt they had (A + SA) improved their knowledge on the course topic by acting as an assessors and providing critical feedback (MV=3.74); 57% (A+ SA) believed they had improved their final version of the task thanks to the peer assessment process (MV=3.5); 82% believed that peer assessment is a valuable learning experience (MV=4.03), although only 35% (A+SA) believed that their peers are capable of providing qualitative feedback to their exercises (MV=3.03) and 35 % (A+ SA) that their peers are capable of assigning a numerical score using the rubric (MV=2.99); 57% (A+ SA) would have liked to assess more than two reports (MV=3.29), and 64% (A+ SA) would have liked to receive more than two reviews (MV=3.39).

Questions	% of students					MV
	SD 1	D 2	N 3	A 4	SA 5	
I used feedback/comments provided by peers to revise the first draft of my task.	2%	20%	20%	47%	11%	3,45
My understanding and knowledge of the topic improved by being an assessor and providing critical feedback	0%	7%	25%	55%	13%	3,74
The quality of my final work improved because of the peer-assessment process.	2%	20%	21%	40%	17%	3,50
Overall, I think that the peer-assessment process was a valuable learning experience.	2%	7%	9%	50%	32%	4,03
I believe my peers are capable of providing qualitative feedback/comments about my exercises.	4%	21%	39%	35%	0%	3,03
I believe my peers are capable of assigning a numerical grade to my task using a standard rubric.	2%	32%	31%	35%	0%	2,99
I would like to have more than two reports assessed.	5%	24%	14%	51%	6%	3,29
I would like to receive more than two reviews.	10%	17%	9%	52%	12%	3,39

SD=Strongly Disagree; D=Disagree; N=Neutral; A=Agree; SA=Strongly Agree; MV=Mean value.

Table 1. Percentage results for a selected subset of the closed questions in the questionnaire.

In Table 2 we show some categories of positive answers to the open-ended questions extracted from the questionnaire, which we qualitatively analyzed for themes regarding common experiences of peer assessment. Column 3 shows the percentage of students who gave a positive answer to each of the listed categories (column 1), which are exemplified by sample comments in column 2.

From the analysis of the closed and open-ended questions we observed that the students' perception about the peer assessment process was very favorable and supports the practice of peer assessment as a valuable learning instrument. The main results of our questionnaire are comparable with similar investigations carried out in different disciplines and at different university levels. For example, Landry et al. (2015) and Amendola and Miceli (2016) also report high appreciation and positive perception of the peer assessment process by their students.

Of the issues reported in the few negative qualitative answers to the questionnaire, the most frequent (25%) are that some students believed the feedback they had received was not detailed enough to help them improve their final work.

To solve this problem, in the next edition of the online Genomics laboratories, we intend to invest more effort in the preparation of a detailed rubric guiding the students to provide more detailed feedback.

Categories of Positive Feedback	Examples of student comments	% of Students
General opinion on peer assessment activity	- I think it is a good way of working. Furthermore, the fact that we have a time table helps us to don't waste time. - According to me the success of the peer assessment depends on the quality of the work of other students and knowledge about this. - Peer Assessment it is a good procedure both for the author and the reader that assessing the peers. It is a very helpful process. - I think that is very important to have a comparison with our peers. Analyzing other works, we can also reflect and improve our task. - I think that the peer assessment is a useful step to improve the reports or to give some cues. - I think is useful to help the students to open their mind to different opinions and to improve their critical ability.	77%
Possibility to improve the report before submitting the final draft	- My ability and possibility to improve my final work is perfect because I was able to understand more of what to do and how to do it. - I think is a good idea, because if you omit something you can add it. - I think it is positive because in this way it is possible to give another look at my work and correct it, especially after the feedbacks by my peers. - Very helpful you get to see other people ideas and compare with your own work. - I think it's a good possibility, because we have more time to overhaul and correct our work in the best way. - I think that is a really positive thing, because in this way everyone has the possibility to do a better work in a correct manner.	73%
Improving critical thinking	-It gave me enough time to prepare and learn by myself, not memorize. -I think this experience improved my critical thinking inducing me to accept the opinion of my peers in order to look at my work in a different way. - My critical thinking is improved thanks to the comparison between my work and those of the other students. - The peer assessment shows me how it is difficult to judge and comment other works. Of course. It would be more difficult judging the works if they were done a little bit better. - After the correction of report of the other student I am able to recognize in a more clear way error of my report. -This experience improved my critical thinking because it abled me to understand different point of view. I could see how many efforts a student can do for a work.	75%
Enhancing learning of the main concepts underpinning the Physics course topics	- I think that even in this case this experience was very useful because I could clarify my doubts so that I was able to enhance my knowledge later. - Peer assessment experience is very help full to understand very well the topics because in this way I can to have a practical approach. - It was a useful experience, in order to write a report it is needed to have a good understanding of the course topics. - It helped me to have a better understanding of about some topics of the genomics course. - With this work I have clarified some doubts and I have learnt concepts about genomics that were not so clear before.	68%

Table 2. Categories of positive qualitative answers extracted from the questionnaire with some examples of student comments[1].

[1] The cited comments are direct quotes from the English responses written by the students.

5. CONCLUSIONS

The main result of this work has been the analysis of the efficacy of the peer assessment activities deployed in the on-line laboratories of Genomics and the response of the students to this collaborative assessment. Different and several data obtained from this experience allowed a detailed investigation of the peer assessment process. The comparison between the teacher's and peers' evaluations, supported by strong statistical correlation of the data, allows us to conclude that the peer evaluation closely correlates with the grades the teacher assigned. Analysing the linear regression between the two variables (peer one grade *vs* peer two grade), we obtained a moderate correlation with a considerable dispersion of the data. This supports the hypothesis of introducing a third peer review, as also requested by the students in the questionnaire.

Our finding of a strong statistical correlation between teacher and peer evaluations is consistent with a previous study (Hongli et al., 2015) which reported a moderately strong Pearson correlation among the variables. These authors also suggested that other conditions can significantly enhance the correlation between teacher and peer evaluations, such as random selection of assessors and assessees, peer assessment being voluntary instead of compulsory, and peer raters being involved in developing the rating criteria. These conditions were adopted in our approach and may be one of the reasons for the stronger correlation found in our work.

Finally, the questionnaire proposed to the students about their beliefs and perceptions about the peer assessment activities allowed us to reflect on many aspects of this process. In general, from the analysis of the answers, we observed that the students' perception is very positive and supports the use of peer assessment as a process to foster students' learning outcomes, knowledge and soft skills, as well as a tool to familiarise students with the dynamics of online collaborative work. These results contribute to confirm that during the collaborative work of peer assessment the students benefited from this intervention. Our work is consistent with conclusions reached in the literature in recent years (Hanrahan & Isaac, 2010; Stevens & Levi, 2005; Nicol et al., 2013, Chao et al., 2010).

Comparison between the grades given by the teacher and the student reports before and after the peer assessment provide evidence that students' learning considerably improved. In addition, our results are in general agreement with previous findings suggesting that peer assessment activities have a differentiated impact on students' learning and perceptions depending on the students' starting level. Moreover, there is also agreement that students across diverse achievement levels generally demonstrate positive perceptions of their peer assessment experience (Lan & Fei, 2015; Li, 2011).

In perspective, quantitative and qualitative analysis of the positive and negative features of the online laboratory permitted us to establish a reference model with guidelines and indicators of quality for future production of collaborative activities in other online laboratories. Indeed, we intend to implement new online laboratories for different scientific disciplines and subjects, featuring enhanced experimental activities and collaborative work. In addition, we will strive to improve the effectiveness of the peer assessment by instructing students to focus their feedback more on suggestions for improvements rather than on numeric evaluation (Nicol et al., 2013).

6. REFERENCES

Airasian, P., & Russell, M. (2008). *Classroom assessment: concepts and applications*. Boston, MA, USA: McGraw-Hill Higher Education.

Amendola, D., & Miceli, C. (2016). Online physics laboratory for university courses. *Journal of E-learning and Knowledge Society*, *12*(3).

Aquario, D., & Grion, V. (2017). Valutazione per l'apprendimento: autovalutazione e valutazione fra pari in alcuni corsi dell'Università di Padova. In E. Felisatti, A. Serbati (Eds.), *Sviluppare la professionalità docente e innovare la didattica universitaria*. Milano, IT: Franco Angeli.

Bangert-Drowns, R., Kulik, C. L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Assessment of Educational research*, *61*, 213-238. doi:10.3102/00346543061002213

Bevitt, S. (2015). Assessment innovation and student experience: a new assessment challenge and call for a multi-perspective approach to assessment research. *Assessment & Evaluation in Higher Education*, *40*(1), 103–119. doi:10.1080/02602938.2014.890170

Chao, I. T., Saj, T., & Hamilton, D. (2010). Using collaborative course development to achieve online course quality standards. *International review of research in open and distance learning*, *11*(3). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/912/1644

Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher*, *1*, 3-31. Retrieved from http://eprints.glos.ac.uk/3609/

Grion, V., Serbati, A., Tino, C., & Nicol, D. (2017). Ripensare la teoria della valutazione e dell'apprendimento all'università: un modello per implementare pratiche di peer review. *Italian Journal of Educational Research*, *19.* Retrieved from http://ojs.pensamultimedia.it/index.php/sird/article/view/2554

Guilford, W. (2001). Teaching peer review and the process of scientific writing. *Advances in Physiology Education*, *25*, 167-175. doi:10.1152/advances.2001.25.3.167

Hafner, J., & Hafner, P. (2010). Quantitative analysis of the rubric as an assessment tool: an empirical study of student peer-group rating. *International Journal of Science Education*, *25*(12), 1509-1528. doi:10.1080/0950069022000038268

Hanrahan, S., & Isaac, G. (2010). Assessing self- and peer-assessment: The students' views. *Higher Education Research & Development, 20*(1), 53-70. doi:10.1080/07294360123776

Hongli, L., Yao, X., Xiaojiao, Z., Mindy, L. K., Youngsun, L., Kyung, S. C., & Hoi, K. S. (2015). Peer assessment in the digital age: a meta-analysis comparing peer and teacher ratings. *Assessment & Evaluation in Higher Education*, *41*(2), 245-264. doi:10.1080/02602938.2014.999746

Kreber, C., Anderson, C., Entwhistle, N., & McArthur, J. (2014). *Advances and innovations in university assessment and feedback*. Edinburgh, UK: *Edinburgh University Press.*

Lan, L., & Fei, G. (2015). The effect of peer assessment on project performance of students at different learning levels. *Assessment & Evaluation in Higher Education*, *41*(6), 885-900. doi:10.1080/02602938.2015.1048185

Landry, A., Jacobs, S., & Newton, G. (2015). Effective use of peer assessment in writing assignment: a case study. *International Journal of Higher Education*, *4*(1), 38-51. doi:10.5430/ijhe.v4n1p38

Li, L. (2011). How do students of diverse achievement levels benefit from peer assessment*?. International Journal for the Scholarship of Teaching and Learning*, *5*(2). doi: 10.20429/ijsotl.2011.050214

Likert, R. (1932). A technique for the measurement of attitudes. *New York archives of Psychology*, *22*(140), 5-55.

Mulder, R. A., Pierce, J.M., & Baik, C. (2014). Peer review in higher education: Student perception before and after participation. *Active Learning in Higher Education*, *15*(2). doi:10.1177/1469787414527391

Nicol, D., Thomson, A., & Breslin, C. (2013). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education*, *39*(1), 102-122. doi: 10.1080/02602938.2013.795518

Persico, D., Pozzi, F., & Sarti, L. (2009). Fostering collaboration in CSCL. *In Encyclopedia of Information Communication Technology* (pp. 335-340). Hershey, NY, USA: IGI Global.

Stevens, D.D., & Levi, A.J. (2005). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback and promote student learning*. Sterling, VA, USA: Stylus.

Topping, K.J., Smith E.F., Swanson, I., & Elliot, A. (2000). Formative peer assessment of academic writing between post graduate students. *Assessment and Evaluation in Higher Education*, *25*(2), 149-169. doi: 10.1080/713611428

Venables, A., & Summit, R. (2003). Enhancing scientific essay writing using peer assessment. *Innovation in Education and Teaching International*, *40*, 281-290. doi:10.1080/1470329032000103816