The Impact of Peer Assessment on the Results of Primary School Students' Self-Assessment in the e-learning System Loomen

Tamara Ređep, Lucija Žignić, Goran Hajdin

University of Zagreb

Faculty of Organization and Informatics

Pavlinska 2, 42000 Varaždin, Croatia

`{tamredjep,lucija.zignic, goran.hajdin}@foi.hr`

Abstract. *The subject of research in this paper is the case study in the field of application of information and communication technology in primary level of education. It deals with understanding the ways in which digital technology can improve the aspects of learning and teaching through an innovative and contemporary way of assessment. The aim of this paper is to apply the formative assessment within the Loomen system in order to investigate the extent to which peer assessment affects students' self-assessment. The results were compared considering deviations from teacher assessment. At the same time, students' satisfaction with such assessment methods was examined. Results indicate that students improved their assessment skills over time and that peer assessment affected their self-assessment. Further, the results indicate that students were consistent in grading through the activities.*

Keywords. Loomen, formative assessment, self-assessment, peer assessment

1 Introduction

Assessment plays an important role in measuring achievement of intended learning outcomes and in providing opportunities for students to self-monitor their learning progress (Jiao, 2015). The basic idea of the new assessment system, implemented as the part of the national curriculum reform is that the evaluation of student achievement should improve the learning and teaching process and in that way students can learn through evaluation. The new system distinguishes three basic approaches to assessment: assessment for learning, assessment as learning and assessment of what has been learned (Bursać et al., 2016).

Assessment of what has been learned is a summative evaluation that involves assessing the level of student achievement after a certain period of learning and teaching during the school year or at its end and usually results with grades. Assessment for learning and as learning are the types of formative assessment, defined as the evaluation of student achievement that takes place during learning and teaching to provide information on student progress. Also its aim is to improve future learning and teaching, encourage student reflections on learning, identify learning deficiencies, identify strength and planning for future learning and teaching. Assessment as learning is based on the methods of self-assessment and peer assessment. In the process of self-assessment, students with the support from a teacher learn to recognize, describe and evaluate their progress in achieving outcomes. Based on this information they direct and adjust their learning and set learning goals. In peer assessment, the student is actively involved in evaluating the learning process and reached achievements of his peers, helping them to observe, monitor and regulate the learning process by giving peer feedback ("Ministry of Science and Education", 2020).

In order to develop lifelong learning skills, it is important that students are engaged in the assessment through peer and self-assessment activities which help students develop self-regulation and self-assessment skills (Ndoye, 2017).

E-assessment is assessment in education employing web tools which deals with the effective use of technology to support successful instruction (Cohen & Sasson, 2016). Digital technology provides opportunities to actively involve students in monitoring and evaluating their own progress. As we live in time when the Internet is easily accessible and where new tools, digital technologies and e-learning systems are being developed daily, digital tools can greatly facilitate access to monitoring and evaluating student achievement and enable their frequent and timely formative assessment. (Tomaš, 2018). Students can benefit from timely and constructive feedback from e-assessment which enables them to evaluate the quality of their work, but this requires higher engagement and substantial involvement in learning

(Jiao, 2015). Therefore, it is important to choose an appropriate digital tool that will motivate students, and positively affect their satisfaction for learning, and provide the support for an active participation in the assessment process.

As the unquestionable importance of involving students in their own learning through self-assessment and peer assessment, the relationship between these two methods is the focal point of study done in this paper. Thus, the main goal of the research is to determine the extent to which peer assessment affects the results of students' self-assessment. Given the aforementioned possibility of digital tools in the application of assessment, connectivity will be explored through the activities within the Moodle based e-learning system called Loomen. Students can access this system with their aai@edu.hr user accounts, and the system itself already has a number of activities that can be used for the purpose of formative assessment, i.e. self-assessment and peer assessment, as well as assessment by teachers.

Based on the defined research goal and the selected research implementation tool, two research questions were formulated:

Q1: *To what extent in the e-learning system does the implementation of peer assessment with accompanying comments affect the results of students' self-assessment?*

Q2: *What do students think about the way of assessment that includes both self-assessment and peer assessment?*

2 Related Work

2.1 Formative e-Assessment

Jiao (2015) in his paper showed results how students can be facilitated in their engagement in learning and how their performances were improved by correcting errors and to receive award marks. Voelkel (2013) described the project which aimed to encourage student's engagement and improvement of their learning by introducing two-stage online tests with a combination of formative and summative approaches to encourage student's engagement. The first stage of each test was formative approach with feedback, and when students achieved 80% they could go to the second summative stage of the test. Described way of combination of formative and summative approaches in the evaluation significantly improved class performance.

Cohen & Sasson (2016) researched the success of students by applying formative evaluation through online quizzes in the Moodle e-learning system. Results indicate that both the average grade on written tests and the average grade on online quizzes were found to be significant predictors of the grade on the final exam and that the students' attitudes towards online quizzes reveals a generally positive attitude. In

their paper Petrović et al., (2017) investigated the effects of using online formative assessments on students' learning achievements. The acquired data suggest that students quickly recognized the value of the formative assessments and that more than 90% of students have used them extensively. Students with available formative assessments achieved significantly higher results in corresponding summative assessments.

In the analyzed papers formative evaluation is conducted with the purpose to stimulate student engagement. The results showed that students who had access to formative assessments improved their performances and their learning and that they can ultimately improve their learning achievements. Results indicate students' positive attitudes about formative assessments.

2.2 Self and Peer Assessment

In their research Romeu Fontanillas et al. (2016) analysed the students' perception about the e-assessment methodology and their role during the process. The results revealed the improvement of the learning process and the high level of satisfaction with the e-assessment activities within the course, especially the activities which focused on self-assessment and peer assessment. Students expressed high satisfaction to the fact of being actively involved in the assessment process. Seifert & Feliks (2019) in their paper also indicated that students noted significant benefit from the assessment process. Peer assessment can help students learn about the assessment process and develop assessment and learning skills. They took responsibility for the assessment, understood its complexity and required level of investment. Anonymous evaluation allowed them to overcome inhibitions in evaluating peers' works as biased, which resulted in improved assessment skills.

Chang et al (2012) provided the comparative analysis in a Web-based portfolio assessment environment for high school students about the difference among teacher-assessment, student self-assessment and peer assessment. The results of self-assessment and teacher-assessment were discovered to be consistent. Oren (2018) also investigated the relationships between these types of evaluations. The results of self-assessment scores and peer assessment scores showed moderately high correlations with teacher scores, but both were higher than the teacher scores. This is confirmed by Šalković et al. (2018), who conducted all three evaluations through the activities of the workshop in Moodle, where regardless of the given elements and evaluation criteria, students evaluate themselves and their colleagues with higher grades than the teacher does.

The results obtained by the qualitative research described by To & Panadero (2019) showed a positive effect of peer assessment on the students' self-assessment process. Peers could aid the self-

assessment process by deepening self-reflection, enriching student understanding of quality and refining subjective judgement. They indicated that peer influence could be reduced by tensions in feedback communication, distrust, competition, and lack of readiness for peer learning.

The results of the analyzed papers indicate high levels of student satisfaction with the self-assessment and peer assessment activities and their involvement in the assessment process. It is also discovered that only the results of self and teacher assessment were consistent and according to the results of self-assessment and peer assessment, students evaluate themselves and their colleagues with higher grades than the teacher does.

3 Methodology

The following variables are defined for the following implementation of the research: a) dependent variable: student self-assessment and b) independent variables: peer assessment, teacher assessment.

The central part of the research is a quasi-experiment which was conducted on primary education level. The quasi-experiment had parallel groups in natural conditions of teaching and learning. The participants were 64 students of the First elementary school in Varaždin. The experiment was conducted in the period of three weeks. Students did not participate in the experiment equally, and only those students who participated in all phases were considered, which ultimately resulted in a sample of 51 students, 22 in the control group and 29 in the experimental group. The sample was structured, based on the class as the smallest unit, and intervention in the class related to the sample was not possible by any of the characteristics. As the influence of class structure is not of primary interest in this research, we can regard it only as a control variable.

Limitation of the conducted research is rather small sample size. Further limitation is the inability to randomly select students within a class, but that is a persistent limitation for research conducted in real life conditions.

The research was conducted as a part of the subject of Informatics among the 7th grade students. As mentioned earlier, the research was conducted through the Loomen system, since the system itself already possesses a number of activities that can be used for the purpose of formative assessment. During the experiment, all work instructions were placed in the newly created course in the Loomen system. The possibility of the Forum was open so that students could receive instructions in a timely manner if they encountered any difficulties.

Within the experiment, three practical tasks were evaluated: I) a Sway presentation, II) a mind map, and III) a digital poster. The assessment criteria was created for each of the tasks, with the number of points for each

of the criteria, and prior to each assessment, all participants in the experiment were introduced to the defined assessment criteria. It is important that students are involved in the self-assessment process and they need to discuss assessment criteria and standards (Angulo & Cebrián-de-la-Serna, 2011). All forms of the assessment were anonymous, in order to minimize potential bias.

In the control group, students performed self-assessment only through the Quiz activity, where they selected one of the offered answers according to the previously defined elements in the assessment criteria. In the experimental group they were acquainted with the fact that in addition to their own assessment, they will be assessed by their peers. The experimental group carried out the assessment through the activity of Workshops, which enables the collection, review, self- and peer- assessment of works through multi-criteria forms. In the peer assessment activity comments were included alongside with the points. After their own assessment and the evaluation of their peers, they saw the grade they received, as well as the accompanying comment. Students then re-performed the self-assessment through the Quiz activity, in order to examine whether there is a difference in self-assessment before and after peer assessment.

Teacher assessment was conducted in all groups in parallel, but independently of self-assessment and peer assessment. Students were familiar with that fact, but during experiment they did not know the grade they got from the teacher. The teacher evaluated each practical work through the activity Assignment based on the same assessment criteria.

At the end of the experiment, students from the experimental group filled out a questionnaire. The questionnaire was completed only by those students who participated in at least one of the activities that included peer assessment, so the number was 38 students, out of which 17 were girls and 21 were boys.

4 Results

4.1 The Results Obtained from the Experiment

Within the experiment, all three forms of assessment were evaluated and compared three times after the completion of the processing of each learning unit. In the experimental group, self-assessment and peer assessment with comments included, were conducted in parallel through Workshop activity. In addition to the evaluation of their work, each student received two works from students in the class. Peer assessment was anonymous and based on points and comments. In order to examine whether the obtained grade and comments affect the student's perception of their own work, students again conducted self-assessment, but this time through Quiz activity. The results were obtained on the change of the deviation between the

self-assessment before and after peer assessment in the experimental group, and then compared to the control group. They were collected from the Loomen system, since statistics were enabled within the system.

Descriptive statistics was conducted to assess the main characteristics of the data. As each of three activities had different maximum grade, the data was standardized to a scale of 0 to 10. Overall grades were high, ranging from minimum of 3,33 to the maximum of 10 overall, 5 to10 when teacher was grading, 3,33 to 10 in peer assessment and 4,44 to10 in self-assessment. In the Table 1., from student's self-assessment average grades by activities and teacher's average grades per activity we can see that on average students' self-assessment was higher than teacher's grade, this goes for all the activities and overall. This is in line with the results from other researchers in the field where the self-assessment grade is usually higher than the teacher's grade (Šalković et al., 2018; Oren, 2018).

If we look at the Table 1 how much the average self-assessment grade per activity differs from teacher's grade, we can conclude that on average, the experimental group had lower discrepancies from the teacher's grade than the control group. Further, we calculated the mean squared error of self-assessment of experimental and control group from the teacher's grade. In the Table 2. it can be seen that the mean squared error (MSE) is lower for the experimental group, therefore meaning lower quadratic deviation of the experimental group from the teacher's grade. Also, it can be noted that experimental group had steeper learning curve then control group, which supports declining MSE for experimental group through activities. To test whether peer assessment has an impact on the grade we used ANOVA with the repeated measures.

Table 1. Arithmetic means of teacher's assessment, student's self-assessment and pear-assessment per activities

	Teacher assess. (Cont.+ exp. gr.)	Self-assess. (Cont. + exp gr.)	Pear assess. (Exp. gr.)
Activity1	8,3987	9,3301	9,1379
Activity2	8,9325	9,4771	9,380
Activity3	9,1993	9,1340	9,3103
SUM	**8,8435**	**9,3137**	**9,2784**

Table 2. Deviations (mean squared error) of student's self-assessment from the teacher's assessment by activities for experimental and control groups

	Activity 1	Activity 2	Activity 3	Overall
Exp. gr.	1,6082	1,3841	1,3490	1,4516
Con. gr.	1,5590	1,4982	1,5691	1,5424
Overall	1,5871	1,4344	1,4481	1,4915

We can see clearly that the structure and the variability of students' grade becomes more similar to teacher's grade through the activities (and time) after the first activity, showing the improvement in the process of learning and the ability of students to grade more objectively (similar to teacher) over time as well as how to take into consideration peer's feedback (Figure 1).

Considering that students' grading skills had improved over time, we were also interested if there was an overall consistency in grading activities or whether it is affecting grades, which we also tested with ANOVA with the repeated measures.
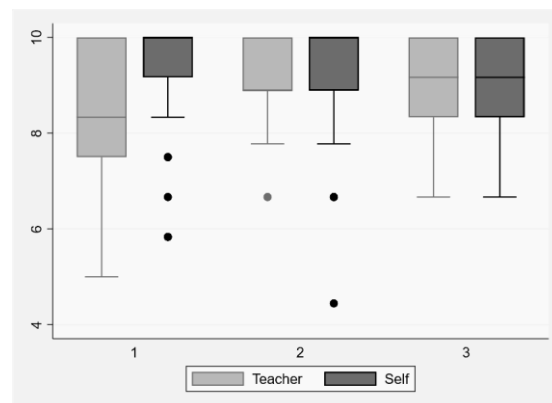

Figure 1. Descriptive statistics - Boxplot diagram of teacher's grade (y axis) and student's self-assessment grade (y axis) per activity (x axis)

For the test purposes the sample of 21 students was taken from the control and experimental group. Then the main assumptions for ANOVA with the repeated measures were checked – within the activities variances and normal assumptions for each activity.

Within subject variances they are similar (Table 3), except for the third activity, on the other hand normality assumption is not met (Figure 2). In the Figure 2 we can see density plot for each activity where x axis represents grades, while y axis represents density. However, practical experiments show that ANOVA is usually robust on normal assumption violations. Finally, the effect of peer assessment was tested with ANOVA with the repeated measures.

Table 3. Variation of student's self-assessment within the group (variances caused by differences within activities)

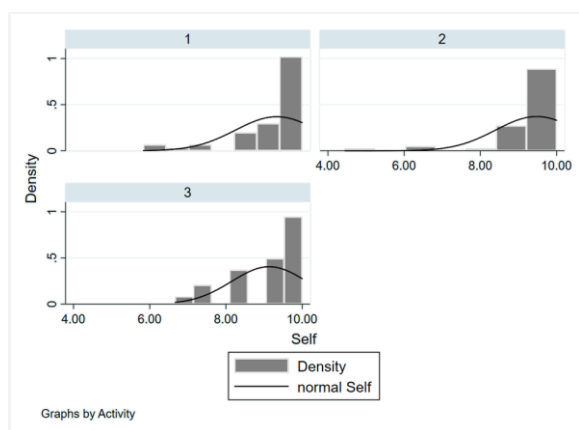|	Activity1	Activity2	Activity3
Activity 1	1,3280		
Activity2	0,6861	1,3392	
Activity3	0,4146	0,0925	0,9082

Figure 2. Histogram and density plot of student's self-assessment of each activity

The obtained results show the significance (P-value of 0,0015) of peer assessment and activity, meaning that peer assessment affects self-assessment. On the other hand, the effect of activity on self-assessment is not significant (P-value of 0,1250). This means that students are consistent in grading through the activities.

Results presented in this chapter provide the answer to Q1: To what extent in the e-learning system does the implementation of peer assessment with accompanying comments affect the results of students' self-assessment? ANOVA results indicated that there is a significant effect of peer assessment to self-assessment. Results show that this effect appears in smaller overall differences of experimental group grades from teacher's grades than it has control group. Other effects can be seen in declining mean square error of student's grades from teacher's grades through activities of the experimental group, while the control group remains constant. This supports that peer assessment helps students to learn how to grade better, meaning more similar to teacher who's grading is used as a benchmark.

4.2 The Results Obtained After the Experiment

The students from the experimental group who participated in at least one of the activities that included peer assessment filled out the questionnaire.

Answering the first question on whether they found the feedback they received about their work useful, 78,95% of student stated that it was useful for them, just 7,89% said that it was not and 10,53 didn't know. If we analyze by gender, then 52,63% of boys and 26,32% of girls said that the information was useful to them, 7,89% of girls and none of the boys thought that feedback has no use for them, and 2,63% and 10,53% did not know the answer to this question (Figure 3). Therefore, we can conclude that most boys (M) and girls (F) find feedback on their work useful.
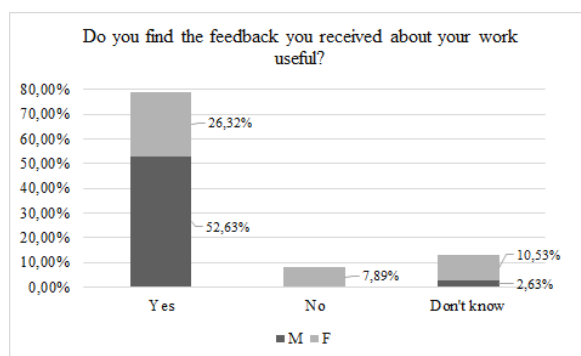

Figure 3. Students' opinion on the usefulness of feedback through peer assessment, n=38

Analyzing the students' responses to the claim that peer assessment influenced their self-assessment, 34,21% of students stated that they were indecisive, and that answer stands out the most on the scale. The answers of the other students were evenly distributed among the other categories. Analyzing by gender, we can see that there is no significant difference between the responses to this statement between boys and girls (Figure 4).
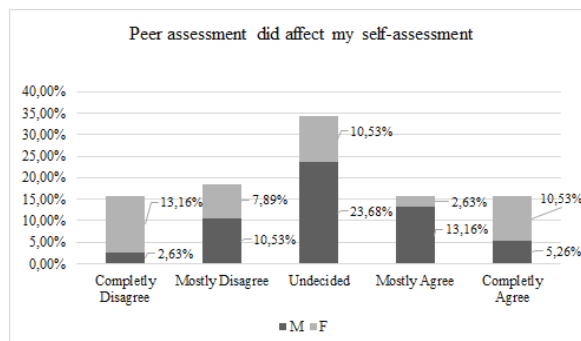

Figure 4. Students' attitudes about the impact of peer assessment on their self-assessment, n=38

Students' opinions on whether they want to use this method of assessment, which includes self-assessment and peer assessment, are almost equal in terms of the number of those who want it and those who do not know if they want it. Thus, 42,11% of students want it, and 47,37% of them do not know if they want it. The equal number of boys and girls do not know if they want to apply this assessment more often, as many as 23,68%, and also equally 5,26% of them do not want it. Among those who want it, we find a smaller difference, because 26,32% of boys and 15,79% of girls want to use this method of assessment (Figure 5).
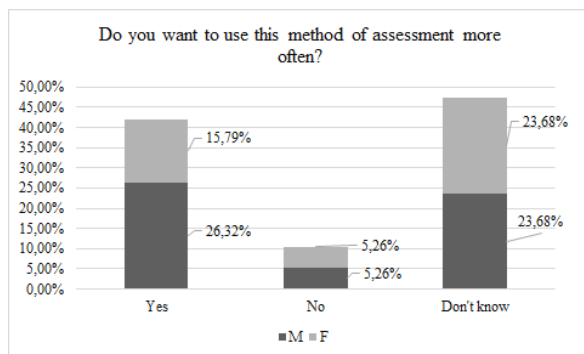

Figure 5. Students' opinion on more frequent usage of self-assessment and peer assessment, n=38

An analysis of responses to claims about how satisfied students feel about the implementation of the assessment within the Loomen e-learning system, shows that the majority of students are completely satisfied, and that claimed 21,05% of boys and an equal number of girls, which is a total of 42,11%. At the same time, an equal number of students who completely disagree with this statement is 2,63% , of a total of 5,26%. Among the answers the largest share (26,38%) was found in boys' statements who mostly agree with the statement that they are satisfied with the use of this assessment in the Loomen system (Figure 6).
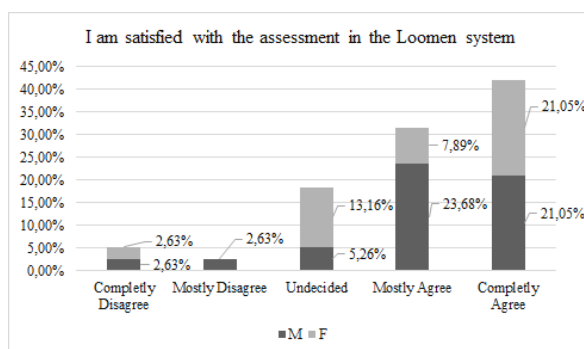

Figure 6. Students' attitudes about the level of satisfaction with assessment in the Loomen system, n=38

Results presented in this chapter provide the answer to the Q2: What do students think about the way of assessment that includes both self-assessment and peer assessment? Majority of the students thought that received feedback was useful for them. On the other hand, most of the students were undecided regarding influence of received peer assessment on their self-assessment. Finally, similar number of students are not sure or think that they do want to use such assessment more often during their learning. Overall, students were satisfied with the assessment implementation in the Loomen.

5 Conclusion

Since educational institutions are increasingly looking for effective and productive technological aids for the development of new models of teaching and learning (Cohen & Sasson, 2016) and also for assessment, the ultimate goal to use digital technology is to create positive patterns for further student development and his access to learning (Tomaš, 2018).

The research area described in this paper is an innovative approach to formative assessment using the Loomen e-learning system. In the given research, the influence of the peer assessment method on students' self-assessment was examined, in comparison with the assessment results obtained by teachers.

The results obtained in the first part of the research provide an answer to the first research question in which we try to examine to what extent in the e-learning system the implementation of peer assessment with accompanying comments affects the results of students' self-assessment. The main result is that peer assessment affects students' self-assessment. We also established that experimental group which had peer-assessment has lower discrepancies from the teacher's grade. However, the obtained results showed that students generally give grades higher than the teacher, but they are consistent in grading.

Answering to the second research question on what students think about the assessment method, which includes both self-assessment and peer assessment, the results indicated that although most students find the feedback they received about their work useful, they are indecisive if peer assessment has an impact on their self-assessment. Although almost half of the students are hesitant in deciding whether they wish to apply this type of assessment which includes self-assessment and peer assessment, they are also completely satisfied with the implementation of the assessment within the Loomen system.

Finally, our research results that peer assessment affects self-assessment and that students find feedback on their work useful, it leads to possible further research directions. Later it could be explored how to conceptualize peer assessment to bring the assessment closer to the one of the teacher.

References

Angulo, J., & Cebrián-de-la-Serna, M. (2011). Study of the impact on student learning using the eRubric tool and peer assessment. In Education in a technological world: Communicating current and emerging research and technological efforts. Edit Formatex Research Center (pp. 421–427).

Bursać, L., Dadić, J., & Kisovar-Ivanda, T. (2016). Učeničkim samovrednovanjem do kvalitetnih učeničkih postignuća. Magistra Iadertina, 11.(1.), 73–88.

Chang, C.-C., Tseng, K.-H., & Lou, S.-J. (2012). A comparative analysis of the consistency and difference among teacher-assessment, student self-assessment and peer-assessment in a Web-based portfolio assessment environment for high school students. Computers & Education, 58(1), 303–320. https://doi.org/10.1016/j.compedu.2011.08.005

Cohen, D., & Sasson, I. (2016). Online quizzes in a virtual learning environment as a tool for formative assessment. Journal of Technology and Science Education, 6(3), 188–208. https://doi.org/10.3926/jotse.217

Jiao, H. (2015). Enhancing students' engagement in learning through a formative e-assessment tool that motivates students to take action on feedback. Australasian Journal of Engineering Education, 20. https://doi.org/10.7158/D13-002.2015.20.1

Ministry of Science and Education. Smjernice za vrednovanje procesa učenja i ostvarenosti ishoda u osnovnoškolskome i srednjoškolskome odgoju i obrazovanju. Retrieved January 24, 2021, from https://mzo.gov.hr/UserDocsImages/dokumenti/Obrazovanje/NacionalniKurikulum/Smjernice/Smjernice%20za%20vrednovanje%20procesa%20ucenja%20i%20ostvarenosti%20ishoda%20u%20osnovnoskolskome%20i%20srednjoskolskome%20odgoju%20i%20obrazovanju.pdf

Ndoye, A. (2017). Peer/Self Assessment and Student Learning. International Journal of Teaching and Learning in Higher Education, 29(2), 255–269.

Oren, F. S. (2018). Self, peer and teacher assessments: what is the level of relationship between them? European Journal of Education Studies, 0(0), Article 0. https://doi.org/10.46827/ejes.v0i0.1661

Petrović, J., Pale, P., & Jeren, B. (2017). Online formative assessments in a digital signal processing course: Effects of feedback type and content difficulty on students learning achievements. Education and Information Technologies, 22. https://doi.org/10.1007/s10639-016-9571-0

Romeu Fontanillas, T., Romero Carbonell, M., & Guitert Catasús, M. (2016). E-assessment process: Giving a voice to online learners. International Journal of Educational Technology in Higher Education, 13(1), 20. https://doi.org/10.1186/s41239-016-0019-9

Šalković, S., Žiljak, V., & Sikirica, N. (2018). Samovrednovanje i ocjenjivanje korištenjem web tehnologija. Polytechnic and design, 6(3), 199–206. https://doi.org/10.19279/TVZ.PD.2018-6-3-09

Seifert, T., & Feliks, O. (2019). Online self-assessment and peer-assessment as a tool to enhance student-teachers' assessment skills. Assessment & Evaluation in Higher Education, 44(2), 169–185. https://doi.org/10.1080/02602938.2018.1487023

To, J., & Panadero, E. (2019). Peer assessment effects on the self-assessment process of first-year undergraduates. Assessment & Evaluation in Higher Education, 44(6), 920–932. https://doi.org/10.1080/02602938.2018.1548559

Tomaš, S. (2018). Prirucnik_Digitalne-tehnlogije-kao-potpora-pracenju-i-vrednovanju.pdf. Retrieved January 18, 2021, from https://pilot.e-skole.hr/wp-content/uploads/2018/03/Prirucnik_Digitalne-tehnlogije-kao-potpora-pracenju-i-vrednovanju.pdf

Voelkel, S. (2013). Combining the formative with the summative: The development of a twostage online test to encourage engagement and provide personal feedback in large classes. Research in Learning Technology, 21. https://doi.org/10.3402/rlt.v21i0.1915