PEOPLE: International Journal of Social Sciences
ISSN 2454-5899

Aleksić & Arula, 2017

Volume 3 Issue 2, pp. 423 - 436

Date of Publication: 08th September, 2017

DOI-https://dx.doi.org/10.20319/pijss.2017.32.423436

This paper can be cited as: Aleksić, K., & Arula, V. (2017). Collaborative Testing – Implications on Knowledge Retention. PEOPLE: International Journal of Social Sciences, 3(2), 423-436.

This work is licensed under the Creative Commons Attribution-Non-commercial 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc/4.0/ or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

COLLABORATIVE TESTING – IMPLICATIONS ON KNOWLEDGE RETENTION

Katarina Aleksić
Branislav Nusić Primary School, Belgrade, Serbia
katarina.aleksic@gmail.com

Verica Arula
Branislav Nusić Primary School, Belgrade, Serbia
verica.arula@gmail.com

Abstract

Long-term monitoring of students' achievements (5th to 8th grade, Branislav Nušić Primary School, Belgrade, Serbia) on Serbian Language and Computer Science tests, led authors to the conclusion that achievements, in most cases, are lower than expected. Test scores are inconsistent with the level of proficiency students are showing at school, during classes, when knowledge is not formally evaluated or is assessed in a different way. It is also noted that the knowledge that students demonstrate on tests usually does not keep. The lack of functional knowledge is confirmed by students' weak achievements on the Primary School Final Exam.

The authors have noticed that process of testing in both subjects causes discomfort in students, a decline of confidence, uncertainty and fear that they will not be successful enough. The authors have searched for a new type of testing that would have a positive impact on students' achievements, prolong knowledge retention and provide the freedom of communication that encourage peer learning and encourage self-testing and auto-correction.

The authors assumed that deficiencies of traditional testing could be overcome by introducing collaborative testing. Through action research, authors have concluded that collaborative testing contributes to the knowledge retention. Students achieve better results on tests. Testing becomes part of the learning process. The students are in a position to assess their knowledge and to act in order to advance. By developing the social aspects of education, progress of other students becomes one's own progress, and responsibility is handed over to the learner at the fullest extent.

Keywords

Collaborative Testing, Knowledge Assessment, Peer Learning, Knowledge Retention

1. Introduction

Branislav Nušić Primary School was founded in 1935 in Belgrade, the capital of Serbia. The school is placed in an urban, multi-ethnic and socially deprived environment. Educational and social background of students' parents[1], as well as the great number of incomplete and dysfunctional families are unfavorable factors for supporting students in learning process. The parents wish their children begin to work as soon as possible. Conclusion is based on the on the analysis of the school choice where students continue their further education (only 15-20% of all students enroll in grammar schools). Such an educational eco-system requires a school practice where strengthening of students' cross-curricular competences is an imperative. To that effect, learning at school is promoted, with a focus on such activities that involve greater cognitive load. Branislav Nušić Primary School is fully oriented towards multidisciplinary, integrative and project-based teaching and learning. The teachers seek to activate the students' potentials through adoption of new, inclusive teaching methods and implementation of ICT in the educational activities.

A long-term monitoring of students' achievements (5^{th} to 8^{th} grade, Branislav Nušić Primary School, Belgrade) on Serbian Language and Computer Science tests, led the authors to the conclusion[2] that the achievements, in most cases, are lower than expected. Test scores are noticeably inconsistent with the level of proficiency that students are showing at school, during classes, when knowledge is not formally evaluated or is assessed in a different way. The

[1] In the generation of 100 students, only 3 parents have University degree.
[2] The conclusion is based on personal professional experience of the authors.

performance-based anxiety, fear of failure and cultural components of fear that impact learning are obvious. It has also been noted that the knowledge demonstrated by students on tests usually does not last long. It may be caused by "short-circuited" key functions of the central nervous system when students are overwhelmed by feelings of fear (Scott Bledsoe and Baskin, 2014). The lack of functional knowledge is also evidenced by students' poor achievements on the Final Primary School Exam.

Having in mind that peer feedback helps student to take more responsibilities in learning process (Khusnia, 2015), the authors have searched for a new type of testing that would have a positive impact on students' achievements, improve the knowledge retention and quality of mutual communication, i.e. interaction that fosters peer learning and encourages self-assessment and auto-correction (Antić, 2010).

2. The Research Problem

By observing the habits and activities of their students, the authors have noticed that students prepare themselves for tests in a short and intense way. Right after the test, students move to the following learning content, while the previous knowledge retains shortly and on a small-scale. No opinion and knowledge exchange has been observed among students, nor does any other form of peer learning that strengthen students' self-confidence regarding learning Serbian Language and Computer Science.

Additional obstacle for the quality and retention of ICT knowledge is the status of Computer Science as a school subject. Computer Science is an elective school subject, unaligned with other optional subjects by nature and complexity of curriculum, with a grade that does not influence the students' overall success. It makes the students less motivated[3], and minimizes their responsibility and chances to master difficult content. During the classes, students show an excellent reasoning and high interest in the topic, however they make no extra effort to learn at home. They do not use available learning support (online school platform and textbooks) and completely rely on the content learned at school.

The authors have noticed that process of testing in both school subjects causes discomfort in students, a decline of confidence, uncertainty and fear that they will not be successful enough.

[3] Personal experience and impression of Katarina Aleksić, Computer Science Teacher.

These shortcomings of traditional testing (hereinafter referred to as **t-test**) could be overcome by introducing **collaborative testing** (hereinafter referred to as **c-test**).

C-test takes place on one school hour (45 minutes), through the following phases:

1. **The student independently completes test (25-30 minutes) and submits it to the teacher.**
2. **The student gets a new test (identical to the previous one), completes it again, now in a group of 3-4 students where the exchange of opinion is encouraged (15-20 minutes), and submits the test to the teacher.**

The teacher evaluates the test by assessing both versions. Score on the first (independent) work makes 70-90% of the final mark, while the score on the second (collaborative) work makes 10-30% of the final mark. Percentage ratio is determined by the teacher, in accordance with the nature of the test, the method of determining collaborative groups and abilities of the students.[4]

According to the research of influence of anxiety and quality of interaction on collaborative test performance, college students with higher levels of test anxiety were most likely to benefit from collaborative testing and to experience the greatest test anxiety reduction. (Pandey and Kapitanoff, 2011). By practicing collaborative testing, the overall mark in the gross anatomy course increased by 9% with a decreased failure rate (Green, Cates, White and Farchione, 2016).

Having in mind that learning is co-construction of knowledge (Pešikan, 2010), and that peer interaction, based on students' equality, is especially beneficial for the development of logic, because it creates conditions for cooperation (Stepanović, 2010), the authors have assumed that c-test has a significant didactic value and that positive effects on learning and knowledge retention can be achieved by its application. Starting from these assumptions, the objective of this study was to determine and compare the effect of t-test and c-test on quality and retention of knowledge and to establish didactic value of the c-test.

3. Pilot Research Study

The pilot research study was conducted in the period from the end December 2014 to the mid-April 2015.

[4] The teacher establishes a high percentage of group work involvement results in the final grade when introducting a new way of testing and when he/she wants to encourage peer learning in heterogeneous groups. This percentage decreases over time in order to strengthen students for high-quality individual work.

3.1 Methodology

The sample in this pilot research study included 43 7th grade students (Branislav Nušić Primary School in Belgrade) – the experimental pilot group (EP-group) which consisted of 21 students and the control pilot group (CP- group) which consisted of 22.

The method utilized was an experiment with parallel groups. Students' knowledge testing regarding both school subjects was conducted according to the following methodology:

Table 1: *Methodology of Pilot Testing*

Experimental pilot group	Control pilot group
1. Traditional test	1. Traditional test
2. Collaborative test (4 weeks after the previous one, announced)	2. Traditional test (4 weeks after the previous one, announced)
3. Traditional test (2 weeks after the previous one, **unannounced**)	3. Traditional test (2 weeks after the previous one, **unannounced**)

All the three tests were to check the knowledge of the same learning contents. After the completion of testing for both school subjects, authors conducted the survey in E-group in order to determine students' attitude towards c-test. The survey included 4 essay-type questions (What is your impression of the c-test? What is the most significant improvement over traditional testing? How did you feel during c-testing? Do you find exchange of opinions with your peers meaningful?). The survey was conducted at school (Arula & Aleksić, 2015).

3.2 Results

The achievements of EP-group and CP-group were assessed through the percentage of correct answers.

Table 2: *Results of Pilot Testing*

Serbian Language		Computer Science	
Experimental pilot	**Control pilot group**	**Experimental pilot**	**Control pilot group**

group (EP-group)		**(CP-group)**		**group (EP-group)**		**(CP-group)**	
Traditional test (unannounced)		**Traditional test** (unannounced)		**Traditional test** (unannounced)		**Traditional test** (unannounced)	
Correct answers	Incorrect answers	Correct answers	Incorrect answers	Correct answers	Incorrect answers	Correct answers	Incorrect answers
53%	47%	36%	64%	55%	45%	45%	55%
Collaborative test (announced)		**Traditional test** (announced)		**Collaborative test** (announced)		**Traditional test** (announced)	
Correct answers	Incorrect answers	Correct answers	Incorrect answers	Correct answers	Incorrect answers	Correct answers	Incorrect answers
52%	48%	31%	69%	65%	35%	41%	59%
Traditional test (unannounced)		**Traditional test** (unannounced)		**Traditional test** (unannounced)		**Traditional test** (unannounced)	
Correct answers	Incorrect answers	Correct answers	Incorrect answers	Correct answers	Incorrect answers	Correct answers	Incorrect answers
59%	41%	28%	72%	**68%**	32%	37%	63%

When comparing the results of the first and second test taken, a drop was noted on the Serbian Language test in the EP-group, although the test had been announced. It supports the thesis that the school content prepared and learned for the first test was quickly forgotten, and that students were not aware that. As for the test in the field of Computer Science, an increase was recorded. Collaborative work was a key to that achievement, considering the fact that the students do not prepare themselves for the Computer Science testing (it is an elective school subject, with a mark that does not influence the students' overall success).

When comparing the first and the third test, a 6% increase in the number of correct answers in EP-group (SerbianL.) and a 13% increase (CompSci.) were recorded, respectively, while the CP-group had an 8% decrease (SerbianL.) and a 7% decrease (CompSci.), respectively.

In the CP-group, knowledge retention was in constant decline, while in the EP-group, after c-test, on the third, unannounced test a 7% increase (SerbianL.) and a 3% increase (CompSci.) in the number of correct answers were recorded, respectively.

Based on the results, we may perceive that c-test contributes to the retention of students' knowledge.

When evaluating c-test, 20 EP-group students (95%) had a positive opinion regarding the new testing methodology. As a significant help, they pointed out the opportunity of retaking the test at the same school class (100%). They emphasized that they felt more relaxed (81%), and that the possibility of exchange of opinions with their peers meant a lot to them (90%). One student (5%) mentioned that he was bothered by the noise during the second part of c-test.

The analysis of the pilot research results undoubtedly leads to the conclusion that collaborative testing has a potential to contribute to better knowledge retention, auto-correction and self-regulation in students and to strengthen them to approach the knowledge testing activity with self-confidence.

4. Continuation of the Research

The results of the pilot research were presented at the XVIII International Scientific Conference "Educational Research and School Practice" – Challenges and Dilemmas of Professional Development of Teachers as Leaders in Education (organized by the Institute for Educational Research, Belgrade, November 2015), and met with a great interest and positive comments. The only concern was related to question of fairness of such knowledge assessment.

The research was continued in 2015 within the program SCOPES [5] (Scientific co-operation between Eastern Europe and Switzerland), in REP-Synergy project: Towards improvement of Research capacities essential for teacher Education and Practices in Serbia and Estonia and is still ongoing.

4.1 Collaborative testing within Serbian Language teaching

Within Serbian Language teaching the research included two 6th grade classes (E1-group and E2-group). The E1-group was characterized by achieving extremely poor results in all school subjects, while the E2-group enjoyed the reputation of being extraordinarily successful. The c-group was a class of peer students having very poor achievements and they had another teacher.

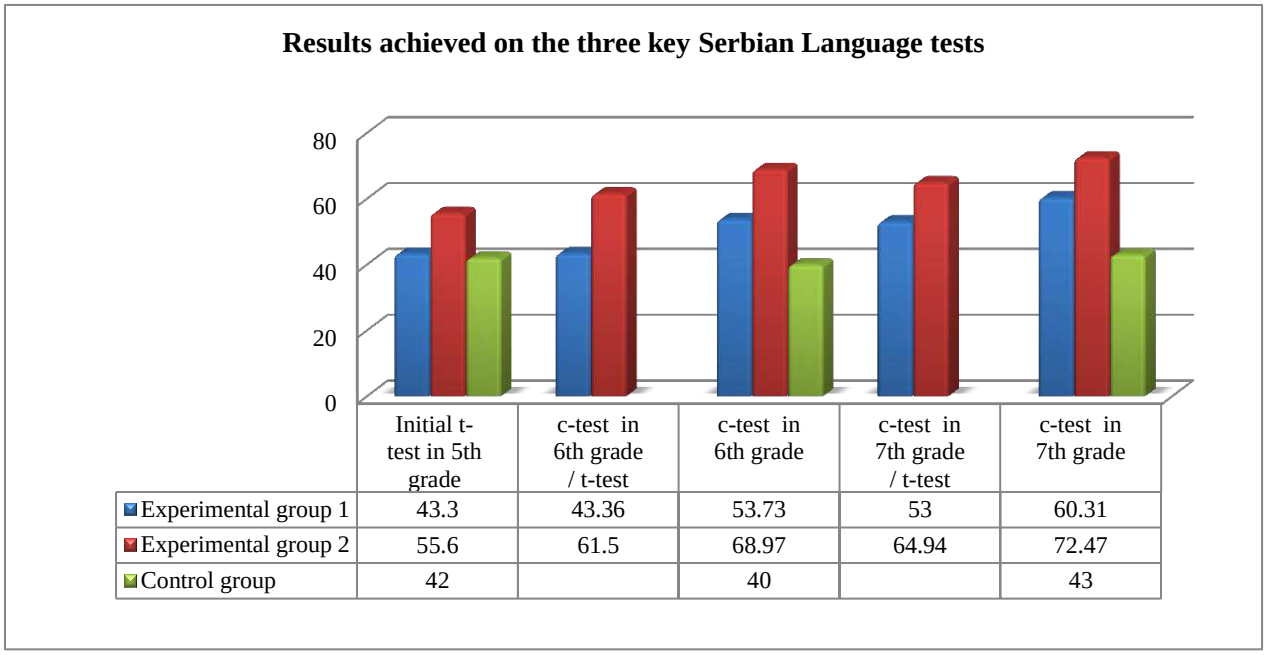
At the beginning of 5th grade (September 2014), the students in all the three groups took the same initial test in the traditional way, and the results of that test were taken as a baseline.

In the following two school years period, the experimental groups were continuously tested collaboratively in accordance with the same methodology as in the pilot research, with the

[5] http://www.snf.ch/en/funding/programmes/scopes/Pages/default.aspx#

control group being always tested in the traditional way. The questions at the tests were the same for all the three groups. All tests were announced.

The results of three key tests (Initial test in 5th grade, Half-term test in 6th grade and Half-term test in 7th grade) are presented in Table 3. Data shows the average number of correct answers when student were working individually, i.e. collaboratively (expressed as a percentage).



	Initial t-test in 5th grade	c-test in 6th grade / t-test	c-test in 6th grade	c-test in 7th grade / t-test	c-test in 7th grade
Experimental group 1	43.3	43.36	53.73	53	60.31
Experimental group 2	55.6	61.5	68.97	64.94	72.47
Control group	42		40		43

Figure 1: *Results achieved on the three key Serbian Language tests*

When comparing the baseline results (t-test) with the results at the half-term test in 6th grade and the half-term test in 7th grade (c-tests), respectively, the experimental groups (regardless of highly different general characteristics) demonstrated a constant increase in the number of correct answers by about 7%, which is somewhat higher than shown by the E-group in the pilot research. This difference could be interpreted as a consequence of a longer time that all these students spent at taking collaborative tests.

The results of these same tests in the C-group indicated stagnation.

When analyzing the results of testing on an individual level, in the experimental groups, the percentage of students who increased their score after collaborative testing is in the range

from 72% to 88%, but the number of those who, based on that, increased their final mark amounts only to about 25%.

Based on these results, we may conclude that the c-test contributes to knowledge retention in the students, and also that it transforms the testing process into a learning situation. Namely, the E1-group had an extremely low level of achievement at the baseline, so that one may suppose that through collaboration the students increased their knowledge level. In the next school year (prior to Final Primary School Exam), overall knowledge retention will be measured.

4.2 Collaborative testing within Computer Science teaching

The curriculum of the Computer Science in Primary School is highly thematically segmented. There is no vertical content integration. The curriculum does not foresee deepening of the content learned in the previous grades.

Based on the positive results of pilot research, a decision was made that during the 2015/16 school year, all 5th grade students take only collaborative test. The basic reason thereof was primarily the need to create an educational environment where the students would learn from each other, co-construct knowledge at school, with an increased extrinsic motivation.

4.2.1 Methodology

The sample included 62 5th grade students who have elected Computer Science course. This age group was chosen, primarily, because of the complexity of the learning contents, and also because of the possibility to implement longitudinal monitoring of knowledge retention.

During the 2015/16 school year, the students were tested 3 times. Test question were related to the topics that are very difficult and abstract for the students' age[6]. In each of the tests, final mark was formed with a different individual and collaborative test parts ratio (Table 4).

Table 4: *Computer Science tests final mark (individual and collaborative work ratio)*

Test 1		**Test 2**		**Test 3**	
Individual	**70%**	Individual	**80%**	Individual	**90%**
Collaborative	**30%**	Collaborative	**20%**	Collaborative	**10%**

The tests were announced, with the key questions being repeated at all the three tests. Collaborative groups were heterogeneous and comprised 3 members each.

[6] Previous long-term experience of Katarina Aleksic, Computer Science Teacher, shows that the students find it very hard to construct knowledge in this field.

In April 2017 (one year after the Test 3) the knowledge retention test was conducted. The test was unannounced. The subject of testing was the content studied in the previous (5^{th}) grade. Certainty-based marking strategy was implemented.

Certainty Based Marking (CBM) is developed by Prof. Gardner-Medwin at UCL (London, UK). It involves asking students not only to answer test question, but also to indicate how certain they are that their answer is correct. After each answer, a student indicates a degree of certainty that the answer will be marked as correct, on a 3-level scale: low, mid or high (Gardner-Medwin and Curtin, 2007).

4.2.2 Results

The students' achievements were assessed through the average number of correct answers expressed in percents (Table 5).

Table 5: *Results achieved by 5^{th} grade students on the three knowledge tests in Computer Science during the 2015/16 school year*

Test 1		**Test 2**		**Test 3**	
Individual work (correct answers)	Collaborative work (correct answers)	Individual work (correct answers)	Collaborative work (correct answers)	Individual work (correct answers)	Collaborative work (correct answers)
41.4%	67.4%	42.8%	74.2%	67.4%	76.7%
<u>By working collaboratively:</u> 76.5% of students increased their score 23.5% of students remained at the same score 0 students decreased their score		<u>By working collaboratively:</u> 60% of students increased their score 26.7% of students remained at the same score 13.3% of students decreased their score		<u>By working collaboratively</u>: 50% of students increased their score 33.3% of students remained at the same score 16.7% of students decreased their score	

35.3% of students got a higher mark **0% of students god a lower mark**	**20% of students got a higher mark** **6.7% of students god a lower mark**	**8.3% of students got a higher mark** **0% of students god a lower mark**

Based on the results shown above, it is clear that collaborative testing helps the students to increase the percentage of correct answers, however it is also clear that small percentage of them also gets a higher mark. Out of 62 students tested on the third test, only three got a better mark.

One year after the Test 3, in April 2017, the knowledge retention test was conducted.

The test was unannounced; it comprised 10 questions, and was conducted as per the traditional method. The contents being the subject of the knowledge retention test had not been mentioned during the 2016/17 school year.

Table 5: *Percentage of correct answers and the level of certainty on the knowledge retention testing conducted in April 2017*

Correct answers (level of certainty)			**Incorrect answers** (level of certainty)		
Low <67%	Mid 67–80%	High >80%	Low <67%	Mid 67–80%	High >80%
3.1%	1.6%	85.9%	1.6%	3.1%	4.7%

Although the contents being the subject of testing had not been studied, nor mentioned during the 2016/17 school year, the students have achieved the results far beyond the expected. With a high level of certainty, 85.9% of the students answered correctly the test questions.

5. Discussion

The results of the pilot research have confirmed that the students, after traditional testing, quickly forget the learned content.

On the other hand, in collaborative testing, the students have enough time to think individually about the test questions, and then a chance to compare their conclusions, through discussion, with the other students' findings. Collaborative testing encourages the students to talk

about the answers, explain their opinion and immediately receive feedback from their peers. This process provides the students with an opportunity to reconsider their standpoint, to make a decision and take responsibility for it. This kind of interaction, primarily because of the feedback, is very important and productive, especially in case of large class size. The c-test adds value to the process of assessment by adding the dimensions of peer teaching, self-correction and self-regulation. It increases the level of quality and knowledge retention. "For, even if overall test performance is not greatly improved by collaborative testing, the positive experiences of increased camaraderie and anxiety reduction it engenders could constitute substantial long-term benefits in the form of increased confidence, motivation, and willingness to continue one's education." (Pandey and Kapitanoff, 2011).

It has been noted that the discussion on the issues and solutions of the collaborative test continued after the test. Consequently, the learning process was significantly intensified and extended.

Continuation of the research has shown that the collaborative testing has much greater effect on the knowledge retention than on the mark itself. We may establish that the mark at the collaborative test is just, since only one of three students with an increased score gets also a higher mark. With the increased frequency of collaborative tests taken, also the students' score increases continuously and the ratio between the score at the individual and collaborative part of the test become ever more balanced. In other words, the difference between the percentage of correct answers when working individually or collaboratively decreases.

Higher score at the test of knowledge is a result of an increased success in the individual part of the test. By practicing collaborative testing, we create an educational situation where collective knowledge becomes also an individual one, so that ever better mark is a result of personal progress of each student.

Collaborative testing has a positive affect both on less successful and more successful students' groups. It has been observed that less successful groups initially achieve large increase in the scores at the collaborative part of the test in comparison with the individual one. There is an evident strong educational impact of several successful students on the rest of the group. The practice of peer support becomes habitual, so that, just several months later, the difference in scores decreases, i.e. the score at the individual part of the test increases.

By implementation of the Certainty-based marking method, authors have got an insight into the level of students' self-confidence in the testing situation. The findings are in favor of

students' high self-confidence when giving correct answers. This method is precious also for the students themselves, because, through feedback, they get a chance to reconsider their own standpoints and to practice auto-correction in learning.

6. Conclusion

Collaborative testing contributes to the retention of knowledge. Students achieve better results on tests. Testing becomes a part of the learning process. A student is in a position to assess his/her knowledge and to act in order to advance. He/she is entrusted with the responsibility for one's own learning to the fullest extent.

Through the process of collaborative testing, a student's individual qualities serve the entire group interests. Also, social aspects of education are developed, where the individual progress has a positive effect on the progress of entire group and vice versa.

On the other hand, while assessing the students' knowledge, the teacher strengthens their self-confidence. By choosing the way to form collaborative groups, and by determining various percentages of proportion of the collaborative mark in the final one, and also by having a perception of the level of students' confidence when taking the test, the teacher has an opportunity to subtly form a student who approaches the testing situation without any stress, and who is capable of demonstrating his/her knowledge.

The research results were presented at the Teachers' Council meeting. This new form of testing has been recognized as a possible solution to the issue of knowledge retention, generally at the level of Branislav Nušić Primary School as a whole. Several teachers of STEM school subjects implement the new method of knowledge testing in students.

7. Scope of Future Research

In the following period, the authors will conduct training for the members of the Teaching School Board (August 2017) on performing and monitoring the collaborative testing.

In the course of the next school year, the authors plan to monitor the impact of collaborative testing on a larger sample of respondents and from multiple teaching areas. The focus of the research will stay on analyzing and measuring the influence by collaborative testing on the knowledge retention through Certainty Based Marking.

References

Arula, V., & Aleksić, K. (2015). The impact of collaborative testing on knowledge retention. XVIII International Scientific Conference "Educational Research and School practice" – Challenges and Dilemmas of Professional Development of Teachers and Leaders in Education: Proceedings. (pp. 286-289). Belgrade: Institute for Educational Research, Institute for the Improvement of Education.

Antić, S. (2010). Kooperativno učenje: modeli, potencijali, ograničenja. Institut za psihologiju Filozofskog fakulteta Univerziteta u Beogradu.

Gardner-Medwin, T., and Curtin, N. (2007). Certainty-Based Marking (CBM) for Reflective Learning and Proper Knowledge Assessment. From the REAP International Online Conference on Assessment Design for Learner Responsibility, 29th-31st May, 2007. Available at http://ewds.strath.ac.uk/REAP07

Green, R. A., Cates, T., White, L. and Farchione, D. (2016), Do collaborative practical tests encourage student-centered active learning of gross anatomy?. American Association of Anatomists, 9: 231–237. https://doi.org/10.1002/ase.1564

Khusnia, A. N. (2015). Strategies to enhance peer feedback and self assessment in extended speaking course. PEOPLE: International Journal of Social Sciences, Vol 1, No 1. https://grdspublishing.org/index.php/people/article/view/344

Pandey, C., and Kapitanoff, S. (2011). The influence of anxiety and quality of interaction on collaborative text performance. Active Learning in Higher Education, Vol 12, Issue 3, pp. 163 – 174. https://doi.org/10.1177/1469787411415077

Pešikan, A. (2010). Savremeni pogled na prirodu školskog učenja i nastave: socio-konstruktivističko gledište i njegove praktične implikacije. Psihološka istraživanja, XIII (2), 157-184.

Scott Bledsoe, T., and Baskin, J. (2014). Recognizing Student Fear: The Elephant in the Classroom. College Teaching, Vol. 62, Iss. 1. https://doi.org/10.1080/87567555.2013.831022

Stepanović, I. (2010). Određivanje vršnjačke interakcije i istraživanje njene uloge u kognitivnom razvoju u kontekstu Pijažeovog i socio-kulturnog pristupa. Psihološka istraživanja, XIII (2), 219-240.