Gamification in Higher Education – a Pilot Study with SQL Course

Georgi Tuparov[1], Daniela Tuparova[2]

[1] *New Bulgarian University, 21 Montevideo str., Sofia, 1618, Bulgaria*
[2] *South-West University "Neofit Rilski", 66 Ivan Mihailov str., Blagoevgrad, 2700, Bulgaria*

Abstract

Nowadays studying Computer Science or Information Technologies is a stable trend for young people who want to have a successful and well-paid career. Unfortunately, part of the students simply follows this trend, instead of being deeply involved in the area of ICT, which calls for additional effort in motivating them. Moreover, the COVID-19 crisis moved a lot of courses into online mode and lecturers had to create or to enrich existing digital learning content and online activities in their courses in order to keep the quality of the educational process. Gamification is possibly one of the modern approaches to increase students' motivation to study more effective.

In this paper, a pilot study of the implementation of gamified self-training and self-assessment in an online SQL course conducted during the spring 2020 and 2021 semesters is presented. Technological and methodological issues of implementing gamification are discussed too. Usage of gamified self-training and self-assessment in SQL course allows higher final achievements of the students regardless of some used gamified elements and number of gamified self-trainings.

Keywords

Gamification, higher education, SQL course, self-training, self-assessment

1 Introduction

The modern information society is immersed in the Internet and many day-to-day activities are performed through or with the help of information and communication technologies (ICT). More and more specialists are needed to support this growing ICT usage, especially in the field of data collection, storage and processing. In this sense, the Databases course is an integral part of Computer Science, Information Systems, Information Technology, and Software Engineering university programs [1]. In addition, since much of the data is organized using the relational model, learning SQL is a prerequisite for building an adequate body of knowledge in the field.

Many young people see a career in ICT as promising and well paid, without having a real idea of the specifics of the field. Moreover, nowadays students prefer to be online all of the time, to game in the virtual worlds and to access learning materials and activities using their mobile devices. Universities and university lecturers are involved in the game to teach these students effectively enough by implementing more ICT in the educational process and using pedagogical approaches appropriate for this online generation. Furthermore, the rising of the COVID-19 crisis has forced the universities to switch into a purely online mode that is in many cases the only solution to conduct classes.

Game-based learning and gamification have also become popular in higher education. Although some authors state that game-based learning is an innovative approach for education at all educational

Education and Research in the Information Society, September 27–28, 2021, Plovdiv, Bulgaria
EMAIL: gtuparov@nbu.bg; ddureva@swu.bg
ORCID: 0000-0003-4162-5106; 0000-0003-0358-0648

levels [2], development of the serious games in higher education is not so easy and gamification is the alternative that can be implemented approximately more easily using e-learning environments. Gamification is "an example of a modern teaching method that can be used in any sphere of human activity, and is applicable to each age group" [3]. S. Deterding at al. define gamification as "the use of game design elements in non-game contexts" [4]. Also, Kapp states in [5]: "Gamification is using game-based mechanics, aesthetics and game thinking to engage people, motivate action, promote learning and solve problems." A set of challenges in front of the university lecturers and students in the implementation of gamification are discussed in [6].

The aim of this paper is to present results from a pilot implementation of gamified self-training combined with self-assessment of students' SQL knowledge and skills in pure online mode during the last two spring semesters conducted in the COVID-19 crisis restrictions.

2 Assessment of students' SQL problem-solving skills in online mode

In order to be effective for students, self-training and self-assessment of their SQL knowledge and skills have to be automated, i.e. students have to receive their assessment results immediately after submission of the solutions. In general, four approaches are suitable for automated online assessment of students' SQL problem-solving skills [7]:

- Creation of a SQL query from scratch to receive a specified result from a database with a given schema;
- Filling the missed part/s in a template of SQL query to receive a specified result from a database with a given schema and optionally state;
- Selecting the appropriate SQL query from the list of queries to receive a specified result from a database with a given schema and optionally state;
- Evaluation of the SQL query result (number of rows returned) from a database with a given schema and state.

2.1 Specific online solutions

One of the first solutions for online evaluation of the created SQL query from scratch is presented in [8]. The authors' idea is to train students in a way that is close to the way professional developers create SQL to solve real tasks – starting from scratch and continuously refining the query after each testing execution up to the moment it becomes syntactically and semantically correct. To implement this idea, an AsseSQL web-based tool for self-training and examination has been developed. The tool provides an online environment for the creation of a SQL query and its execution against an example database in some states. During the process of SQL query creation, AsseSQL simulates the traditional command-line SQL interface to the relational database so students can run the SQL query they have created as many times as needed, and receive feedback about its syntactical and semantical correctness. To avoid the possibilities to submit a query that returns the desired result, but only in the current state of the example database, the query is tested for correctness with a different hidden database state, which is not available to students [9]. No gamification options are available in this tool.

Some other tools provide similar functionalities – SQLator [10], testSQL [11], and SQL Tester [12]. With these tools, students are able to self-train creation of single table queries (SELECT-FROM-WHERE) optionally with GROUP BY and HAVING clauses, two tables join and/or simple subqueries. Students' achievements can be used in the process of self-training or official grading. Gamification functionalities are very limited or missing at all.

HackerRank [13] is a good and very popular environment for self-training and self-assessment in programming languages including SQL. After free registration, students can start creating SQL queries to solve problems and receive points for every correct solution. Some gamification elements are available like moving through the levels, receiving awards (stars) and leaderboard ranking. The lecturers could manage the self-training and assessment process creating contests for the students.

Some online games for the creation of SQL queries from scratch are available for free – SchemaVerse [14], SQL Island [15], and GalaXQL [16]. Unfortunately, the tasks included in these games are fixed and cannot be changed according to the learning goals of the particular syllabus.

The second and third assessment approaches – filling the missed part/s in a template of SQL query or selecting the appropriate SQL query from the list of queries in order to receive a specified result from a database with a given schema and optionally state, are widely implemented in many web sites for self-training and self-testing such as w3schools.com [17], sqlquiz.com [18], etc.

The fourth assessment approach – by evaluation of the SQL query result as a number of rows returned from a database with a given schema and state is not popular in the online systems for self-training and self- assessment, probably because it needs a database schema with a state-provided [7].

All of the systems mentioned above are not compatible with the standards for exchanging learning content or information about students' achievements among e-learning environments like LTI [19], xAPI [20], and CMI5 [21], so they cannot be integrated into the universities' e-learning ecosystems.

2.2 Solution in e-learning environment

Unfortunately, assessing open answer questions is not well enough implemented in nowadays e-learning environments so automated evaluation of the created SQL queries is not applicable without the creation of the extra modules or services. On the other hand, filling gaps in the template, multiple-choice and short answer questions are very well implemented and can be used for implementation of the last three approaches for self-training and self-assessment mentioned before. In pure online mode, the database schema and state have to be provided with the task.

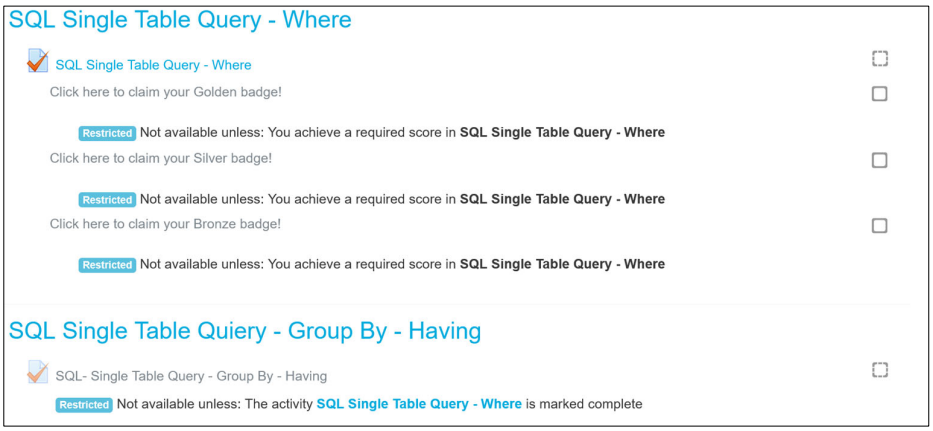

Figure 1: Students' view of self-training tests with badges (Moodle LMS)

Name	Number of badges from this course owned by this student
Someone	12
Someone	12
Someone	11
Someone	10
Someone	10
Student NN	10
Someone	9

Figure 2: Student's view of the leaderboard (Moodle LMS)

For the pilot testing, the fourth testing approach was selected – evaluation of the SQL query result (number of rows returned) from a database with a given schema and state. An implementation as short-answer questions has been done in Moodle. Gamification is organized with the use of badges (bronze, silver or golden) which students can obtain by achieving an appropriate level of correct answers (50%,

70% or 90% respectively). Also, the next self-training test becomes available after passing the current test with at least 50% correct answers (Figure 1).

The students can compare anonymously their achievements with other students (in terms of badges obtained) in a leaderboard (Figure 2).

To increase the productivity of creation of the training and exam tasks, a web-based system for automatic generation of SQL tasks was developed [7, 22] with the following functionalities:

- Storing and managing relational database schemas and states;
- Storing SQL query sets related to the stored database schemas;
- Automatic execution of stored query sets using related database schema in different states;
- Converting the SQL query code to an image file before including it into a generated question.
- Generation of the short-answer questions in Moodle XML [23] and in IMS QTI 2.1 [24] formats for uploading in Moodle or other compatible e-learning systems.

Unfortunately, the system is not able to automatically include images for the database state, so database images are included manually after the generation of the test tasks. An example of the question tasks imported in Moodle is presented in Figure 3.

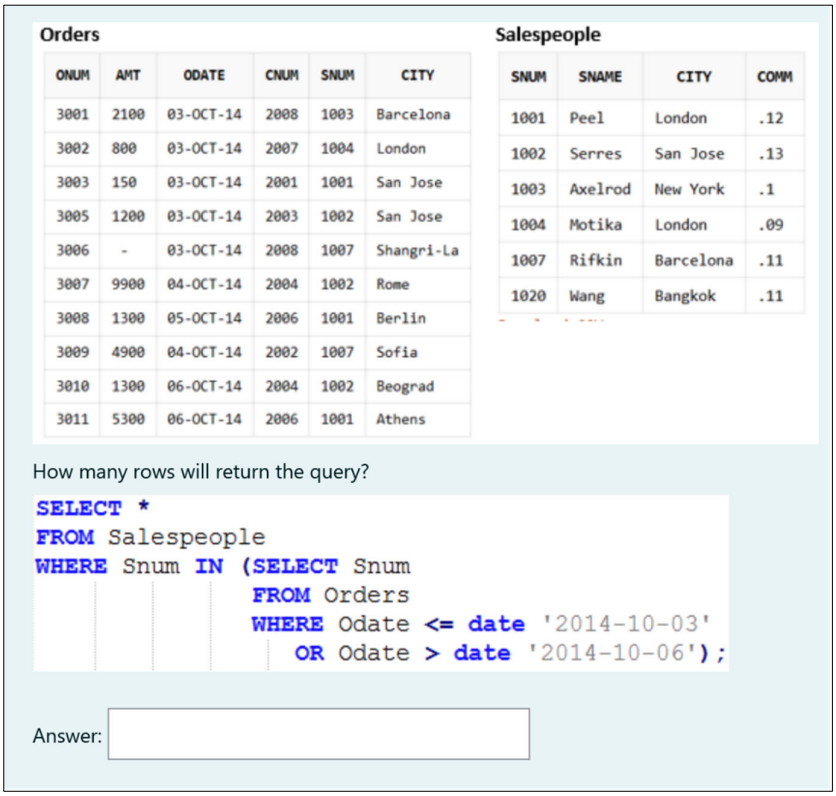

Figure 3: SQL Task with DB schema fragment included (Moodle LMS)

2.3 Pilot testing

Four training tests are set that cover queries on a single table with WHERE clause, a single table with data grouping, queries on more than one table using JOIN, subqueries or set operators. The tests are arranged sequentially and gamification with level achievements is used – in order to move on the next, the student must have solved the previous one with at least 50% success. The number of attempts is not limited. In addition, during the spring 2020 semester additional gamification elements – badges were incorporated. After reaching 50%, 70% or 90% the student can get a bronze, silver or gold badge, respectively. An anonymous ranking of the obtained badges is available, and the student can see his/her achievement compared to the achievements of the other students without seeing their names.

3 Target groups

In the 2020 year 34 students participated, 17 (50%) of them followed gamification with leaderboard, badges, and level achievements. During the 2021 academic year in the pilot testing of the gamified SQL course, 63 students participated – about half of them took part in the gamified training only with level achievements and did not wish to take badges or to follow the leaderboard.

For level 1 the students had to make two trainings and self-assessment (Tr1 and Tr2) and to pass finaltest1 (FinalT1), and for level 2 they also had to make two self-trainings (Tr3 and Tr4) and to pass finaltest2 (FinalT2). The final result (Final_all) is the average of FinalT1 and FinalT2.

For the analysis of the results, we have divided the students into four groups – students who participated in a gamified manner in 2020 – Gam20, and students who did not – nonGam20; students who participated in the second type of gamification in 2021 – Gam21, and students that did not – nonGam21.

The students are grouped in several variables: Grouping_all – with groups Gam20, Gam21, nonGam20, nonGam21; Number of trainings – from 0 to 4 conducted by the student.

4 Data analysis

4.1 Statistical methods

All used statistical methods are applied with significance level α=0.05.

We applied the following statistical methods:

- Shapiro-Wilcoxon and Kolmogorov-Smirnov tests for estimation of normality of the distributions with H0 – The sample is normally distributed, and with the alternative hypothesis – The distribution is not normal;
- Chi-Squared Tests for ratio estimation;
- Descriptive statistics;
- Kruskal-Wallis test for k-independent samples with H0 – The samples come from similar distributions, and alternative hypothesis – The samples come from different distributions;
- Man-Whitney U test for two independent samples with relevant effect size and H0 – the samples have similar distributions; the alternative hypothesis is according to descriptive statistics of the samples.

Some of the samples are not normally distributed and the number of observations in some samples is less than 30 therefore we use non-parametric statistical methods for independent samples.

The data is processed with JASP statistical software, MS Excel and IBM SPSS.

4.2 Research questions

Our research questions are:

- **RQ1.** Are there differences in students' achievements between groups that participated in gamified self-training and those that did not participate?
- **RQ2.** Are there differences in students' achievements in both groups that participated in gamified self-trainings?
- **RQ3.** Did the number of passed self-trainings influence the students' achievements?

4.3 Results

The descriptive statistics of the students' achievements in Final_all are presented in Table 1. The box-plot diagram (Figure 4) shows the differences between gamification and non-gamification groups.

Table 1: Descriptive Statistics in distributions of Final_all Variable in Grouping_all

	Final_all			
	Gam20	Gam21	nonGam20	nonGam21
Valid	17	31	17	32
Missing	0	0	0	0
Mean	83.407	86.290	48.358	58.294
Std. Error of Mean	2.863	2.678	6.343	5.079
Median	84.167	90.833	54.167	61.875
Std. Deviation	11.803	14.910	26.155	28.733
Minimum	58.750	24.583	4.167	0.000
Maximum	97.917	100.000	88.750	95.833

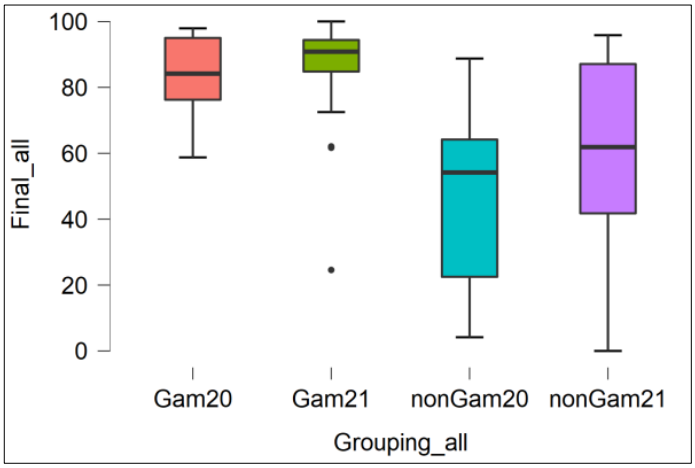

Figure 4: Box-plots for variable Final_all according to groups in variable Grouping_all

Descriptive statistics in groups regarding the number of conducted trainings are presented in Table 2 and Figure 5.

Table 2: Descriptive Statistics of Final achievements according to number of gamified self-trainings

	Final_all				
	0	**1**	**2**	**3**	**4**
Valid	49	3	11	4	30
Missing	0	0	0	0	0
Mean	54.847	89.583	77.576	84.375	87.778
Std. Error of Mean	4.000	2.836	6.128	1.497	2.113
Median	54.583	90.833	82.917	85.417	91.458
Std. Deviation	28.001	4.912	20.323	2.995	11.575
Minimum	0.000	84.167	24.583	80.000	58.750
Maximum	95.833	93.750	97.917	86.667	100.000

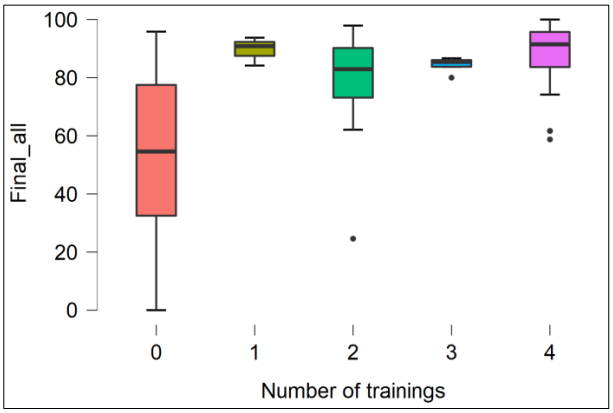

Figure 5: Box-plots for variable Final_all according to groups in variable Number of trainings

Check for normality of distributions

Tests of Normality of distributions of achievements in groups Gam20, Gam21, nonGam20, nonGam21 are performed as shown in Table 3. For group Gam21 we have to reject the null hypothesis about the normality of distribution and accept the alternative hypothesis – the sample does not come from Normal distribution. For the other groups, we can accept the null hypothesis.

Table 3: Tests of Normality in groups Gam20, Gam21, nonGam20, nonGam21

		Kolmogorov-Smirnov[a]			Shapiro-Wilcoxon		
	Grouping_all_T	Statistic	df	Sig.	Statistic	df	Sig.
Final_all	Gam20	.131	17	.200*	.925	17	.176
	Gam21	.229	31	.000	.707	31	.000
	nonGam20	.176	17	.168	.945	17	.383
	nonGam21	.123	32	.200*	.935	32	.054

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Regarding tests of Normality in the groups according to the number of conducted gamified self-trainings we can accept the null hypothesis for normality of distributions of achievements in groups of 0, 2 and 4 self-trainings conducted. For groups 1 and 3 self-trainings the number of observations is too small. (Table 4)

Table 4: Tests of Normality of achievements in groups according to the number of gamified self-trainings.

		Kolmogorov-Smirnov[a]			Shapiro-Wilcoxon		
	Number of trainings	Statistic	df	Sig.	Statistic	df	Sig.
Final_all	0	.107	49	.200*	.951	49	.041
	1	.267	3	.	.951	3	.576
	2	.232	11	.100	.800	11	.010
	3	.333	4	.	.828	4	.163
	4	.214	30	.001	.829	30	.000

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

RQ1. Are there differences in students' achievements between groups that participated in gamified self-training and those that did not participate?

There are significant differences in the achievements in groups that conducted gamified self-trainings and groups that did not (Table 5) ($p<.0001$ with large effect size $r>0.5$ (https://www.spss-tutorials.com/effect-size/)).

Table 5: Independent Samples Man-Whitney U-Test.

	Samples	W	df	p	Rank-Biserial Correlation
Final_all	Gam20-nonGam20 [1]	254.000		< .001	0.758
	Gam21-nonGam21 [2]	786.500		< .001	0.586

Note. For the Mann-Whitney test, the effect size is given by the rank biserial correlation.

Note. [1] For all tests, the alternative hypothesis specifies that group *Gam20* is greater than group *nonGam20*.

Note. [2] For all tests, the alternative hypothesis specifies that group *Gam21* is greater than group *nonGam21*.

RQ2. Are there differences in students achievements in both groups that participated in gamified self-trainings?

In the pilot study, we found that the final results do not show a statistically significant difference in the students' achievements. Mann-Whitney U test is applied with alternative hypothesis group *Gam20* is less than group *Gam21*. (Table 6)

Table 6: Mann-Whitney U test for groups Gam20 and Gam21.

	W	df	p	Rank-Biserial Correlation
Final_all	203.500		0.100	-0.228

Note. For the Mann-Whitney test, the effect size is given by the rank biserial correlation.

Note. For all tests, the alternative hypothesis specifies that group *Gam20* is less than group *Gam21*.

Although we applied different gamification elements in both groups, there is no statistically significant difference between students' achievements.

RQ3. Did the number of conducted self-trainings influence the students' achievements?

Kruskal-Wallis test for comparison of k-independent samples suggests that there is a statistically significant difference ($\chi 2=33,650$ (3), $p=0.000<0.05$).

There is no statistically significant difference in students' achievements in groups of students that conducted self-training regardless of the number of trainings – 1, 2, 3 or 4 (Kruskal-Wallis test, $\chi^2=35.293(3)$, $p=0.152>0.05$). However, a difference occurs between students that have not used self-trainings and students that have, regardless of the number of self-trainings (Table 7).

Table 7: Test Statistics Mann-Whitney U test for groups 0 (without self-training) and conducted the number of self-trainings.

	Final_all (0-1)	Final_all (0-2)	Final_all (0-3)	Final_all (0-4)
Mann-Whitney U	21.000	144.000	44.000	203.000
Wilcoxon W	1246.000	1369.000	1269.000	1428.000
Z	-2.061	-2.398	-1.819	-5.375

Asymp. Sig. (2-tailed)	.039	.016	.069	.000
Exact Sig. [2*(1-tailed Sig.)]	.037[b]		.071[b]	
Exact Sig. (2-tailed)	.036	.015	.069	.000
Exact Sig. (1-tailed)	.018	.008	.035	.000
Point Probability	.002	.000	.001	.000

a. Grouping Variable: Number of trainings

b. Not corrected for ties.

We can reject H0 that there is no statistically significant difference between achievements of students that did not participate in gamified self-trainings and those that participated, and accept that a course with gamified self-trainings, regardless of their number, provide higher students' achievements than a course without such.

5 Conclusions

The gamified pure online self-training and self-assessment of SQL problem-solving skills were tested in the spring semester of the 2020 and 2021 academic years with a total of 97 students from the Informatics Department of the New Bulgarian University. More than half of the students in the courses took advantage of the opportunity for self-training and self-evaluation. In general, the feedback is positive – students like being able to prepare in conditions as close as possible to the exams and being able to compare their achievements with other students in the course (in the spring 2020 semester). The usage of gamified self-training in SQL courses gives the opportunity for higher final achievements of the students regardless of some used gamified elements and the number of gamified self-trainings. Hopefully, the discussed approach for gamified online self-training will be a useful example for university lecturers on how to motivate students to attend pure online activities in the educational process.

Acknowledgements

The presented study is partially supported by the Scientific Fund at the South-West University "Neofit Rilski" Blagoevgrad, Bulgaria and the project DN 05/10 2016 Pedagogical and technological issues of game-based learning, funded by Bulgarian National Science Fund.

References

[1] ACM, ACM Curricula Recommendations, URL https://www.acm.org/education/curricula-recommendations, ACM, 2013, last accessed on 20.08.2021.

[2] Terzieva, T., Golev, A., & Stavrev, C., Serious games – innovative tool for education. In proc. Innovative software tools and technologies with application in mathematics, informatics, and education. 2017, pp. 104-107, University of Plovdiv "Paisii Hilendarski"

[3] Stoyanova, M., Tuparova, D., Samardzhiev, K.Impact of motivation, gamification and learning style on students' interest in maths classes – a study in 11 high school grade, Advances in Intelligent Systems and Computing, 716, pp. 133-142., 2018, DOI: 10.1007/978-3-319-73204-6_17, Springer Verlag, ISSN: 21945357, ISBN: 9783319732039.

[4] Deterding, S., Dixon, D., Khaled, R., Nacke, L., From game design elements to gamefulness: Defininggamification, in: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, ACM, 2011, pp.9–15. http://dx.doi.org/10.1145/2181037.2181040.

[5] Kapp, K., The Gamification of Learning and Instruction: Game-based Methods and Strategies for Training and Education. New York: Pfeiffer, 2012, ISBN: 978-1-118-09634-5

[6] Mora, A., Riera, D., Gonzalez, C., & Arnedo-Moreno, J. Gamification: a systematic review of design frameworks. Journal of Computing in Higher Education, 29(3), 2017, 516–548. doi: https://doi.org/10.1007/s12528-017-9150-4.

[7] Tuparov, G., Stanchev, M., Assessing students' SQL knowledge and skills in an e-learning environment, Computer Science and Education in Computer Science, vol. 1, 2019, pp. 77-79.

[8] Prior, J. C., Lister, R., The Backwash Effect on SQL Skills Grading, ITiCSE '04, June 28-30, 2004, Leeds, United Kingdom.

[9] Prior J. C., AsseSQL: an Online, Browser-based, SQL Skills Assessment Tool, ITICSE'14, June 21–25, 2014, Uppsala, Sweden.

[10] Sadiq, S., M. Orlowska, Sadiq, W., Lin, J., SQLator—an online SQL learning work-bench. In: Proceedings of ITiCSE'04, pages 223–227, Leeds, UK, June 2004.

[11] License, J., testSQL: Learn SQL the Interactive Way. In Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education – ITiCSE'17. ACM Press, New York, USA, 376–376.

[12] Kleerekoper, A., Schofield, A., SQL Tester: An Online SQL Assessment Tool and Its Impact, TiCSE'18, July 2–4, 2018, Larnaca, Cyprus.

[13] HackerRank, URL www.hackerrank.com, last visited on 21.08.2021.

[14] SchemaVerse, URL https://schemaverse.com/, last visited on 21.08.2021.

[15] SQL Island, URL https://sql-island.informatik.uni-kl.de/, last visited on 21.08.2021.

[16] GalaXQL, URL http://sol.gfxile.net/g3/#, last visited on 21.08.2021.

[17] w3schools.com, URL https://www.w3schools.com/quiztest/quiztest.asp?qtest=SQL, last visited on 21.08.2021.

[18] SQL Quiz, URL http://www.sqlquiz.com/sql-quiz.php, last visited on 22.08.2021.

[19] IMS Global Learning Consortium, Learning Tools Interoperability Core Specification, URL http://www.imsglobal.org/spec/lti/v1p3/, last visited on 22.08.2021.

[20] Advanced Distributed Learning Initiative, xAPI Technical Specifications, URL https://adlnet.gov/projects/xapi-technical-specifications/, last visited on 22.08.2021.

[21] AICC, CMI5 Specification, URL https://aicc.github.io/CMI-5_Spec_Current/, last visited on 22.08.2021.

[22] Tuparov, G., Keremedchiev, D., Assessing students' SQL knowledge and skills in gamification manner, MIPRO 2020 – Proceedings, art. no. 9245179, pp. 1531-1536. DOI: 10.23919/MIPRO48935.2020.9245179, IEEE, ISBN: 9789532330991

[23] Moodle XML format, URL https://docs.moodle.org/311/en/Moodle_XML_format, last visited on 22.08.2021.

[24] IMS Question & Test Interoperability (QTI) Specification, http://www.imsglobal.org/question/index.html, last visited on 22.08.2021