Learner Attitude, Educational Background and Gender Influence on Knowledge Gain in a Serious Games-Enhanced Programming Course

Dan Zhao, Cristina Hava Muntean, Adriana E. Chis and Gabriel-Miro Muntean, *Senior Member, IEEE*

Abstract

Contribution: This research study deploys three serious games with various topics in an entry-level C Programming module, and investigates students' learning outcomes. The study also explores whether learners belonging to different subgroups benefit more from the use of serious games than their peers. The subgroups are formed based on learner demographics that capture learners' attitude towards school or STEM subjects, their previous educational performance and gender.

Background: Despite the latest rapid economy growth in the ICT sector, many European countries are facing the challenge of retaining students in STEM related subjects, which could lead to unfilled vacancies in the ICT job markets in the near future. Serious games have been utilized in the classrooms of many STEM subjects to improve students' learning experience and learning outcomes, and potentially encourage their engagement with STEM related industries. While some prior works had assessed the effectiveness of serious games in improving students' learning outcomes, little research has been done to investigate the impacts among students with different previous educational background and performance, attitude, and gender.

Research Questions: 1) Do the proposed serious games improve students' learning outcomes?, and 2) Do students with certain previous educational performance, learning attitude, and/or gender benefit more than others from the use of serious games?

Methodology: To thoroughly study students' learning outcome, a large scale pilot was deployed as part of the first-year undergraduate C Programming module at Dublin City University, Ireland. A multi-dimensional pedagogical assessment toolkit was utilized. In particular, a demographic questionnaire was carried out before the pilot began, based on which students were divided into different subgroups in terms of educational ability, initial attitude to school, attitude towards learning STEM subjects, and gender. Pre- and post-tests were conducted right before and after playing each game. For students belonging to each subgroup and for all students, the average pre- and post-test marks related to each game were compared, while paired-sample t-tests were also conducted to assess the statistical significance of knowledge gain.

Findings: Statistically significant knowledge gains were observed in all three games for all students. The students with good previous educational performance or strong attitude towards attending school and learning STEM subjects, tend to gain more by using serious games, as they obtained statistically significant improved learning outcomes in all three games. The students with average starting points, although performed better in the post-test than in the pre-test in all games, they achieved statistically significant improvements in some cases only.

Index Terms

STEM education, technology enhanced learning, serious games, programming, knowledge assessment.

I. Introduction

In this paper, is investigated the impact of serious games on undergraduate students' learning outcome when learning Programming. Three serious games for an entry-level C Programming module were proposed and developed as part of the EU-funded NEWTON Project[1], covering the topics of *variables*, *loop*, and *structure*, respectively. The paper describes the results of a 12 week pilot deployed at Dublin City University, Ireland, in which the proposed serious games were used in a software development module by first-year undergraduate students with little or no prior experience of programming. The impact of the proposed games on students' learning outcomes was studied. In particular, factors such as students' previous educational performance, students' attitude towards attending school and learning STEM, and gender, were taken into consideration when analyzing the impact.

The paper is structured as follows: Section II presents current research works related to using serious games in education; Section III presents an overview of the NEWTON project, and provides the details of the proposed Programming serious games; Section IV describes the overall setup and participants in this large scale pilot, as well as the evaluation methodologies; Section V presents the results of the study, assessing the learning outcomes achieved through the proposed serious games among all students as well as among students belonging to different subgroups; Section VI and Section VII discuss the results and conclude the paper.

Dan Zhao and Gabriel-Miro Muntean are with School of Electronic Engineering, Dublin City University, Ireland (email: gabriel.muntean@dcu.ie).

Adriana E. Chis and Cristina Hava Muntean are with School of Computing, National College of Ireland, (e-mail addresses: adriana.chis@ncirl.ie, and cristina.muntean@ncirl.ie).

[1]EU NEWTON Project Website [Online] Available: http://newtonproject.eu

II. Related Works

A serious game is a game that has an educational or instructional purpose with well-defined objectives, in terms of transmitting knowledge or supporting skill acquisition [1]. Based on their scope, serious games are used in a wide range of fields including education, healthcare, military, politics, and scientific research [2], [3]. This paper focuses on relevant research employing serious games in education (also referred to as game-based learning (GBL)) to improve students' learning experience and their performance. There is a significant body of research in this area, that has indicated that serious games are effective [4], [5], especially when teaching difficult subjects such as science, technology, engineering, and mathematics (STEM) [6], [7], [8].

Previous studies show that game-based learning foster student motivation [9], [10], engagement, self-efficacy, and overall increase learning experience [6]. However, the games have to be well designed and at the right level of complexity as learners who are bored by the games develop superficial problem solving skills only [9]. Similar findings are reported by Hamari *et al.* [6] who highlight that the challenge of the game directly influences the perceived learning. It was observed that a learner loses interest in playing the game when the game has a low challenge. In contrast, an increase level of challenge, which matches the learner's competence, activates learner's engagement with the game activities which in turns enhances learning by supporting knowledge development.

Serious games have the potential to promote active and self-directed learning through their entertaining and engaging characteristics, enabling students to learn while exploring concepts in authentic contexts and receiving immediate feedback.

In particular, serious games are effective in teaching computer programming as they help students to visualize abstract and difficult programming concepts [11], [12], [13], [14], [15], [16], [17]. Well designed educational games support learners to enhance their computational thinking skills while playing the games, and aid learners to understand fundamental programming constructs [18]. For example, Eagle and Barnes, present Wu's Castle [11], an interactive educational game to support learning loops and arrays. Tessler *et al.* [13] use game-based learning to help students understand the concept of recursion. Miljanovic and Bradbury propose RoboBug [19], a puzzle-based serious game, to support students to learn different debugging techniques to identify bugs in their programs. Zhao and Muntean propose the interactive Warehouse game [16]to learn concepts of variables and data types. Dicheva and Hodge present an educational game [20] that helps students understand and implement the stack data structure.

Despite the popularity of game-based learning, there is no consistent methodology for evaluating the usefulness of a serious game. Different studies may explore one or more dimensions to assess the effectiveness of a game. The dimensions/metrics commonly evaluated are learner engagement [12], enjoyment to learn with the game [14], [18], [19], and attainment of learning outcomes [11], [14], [19], [21].

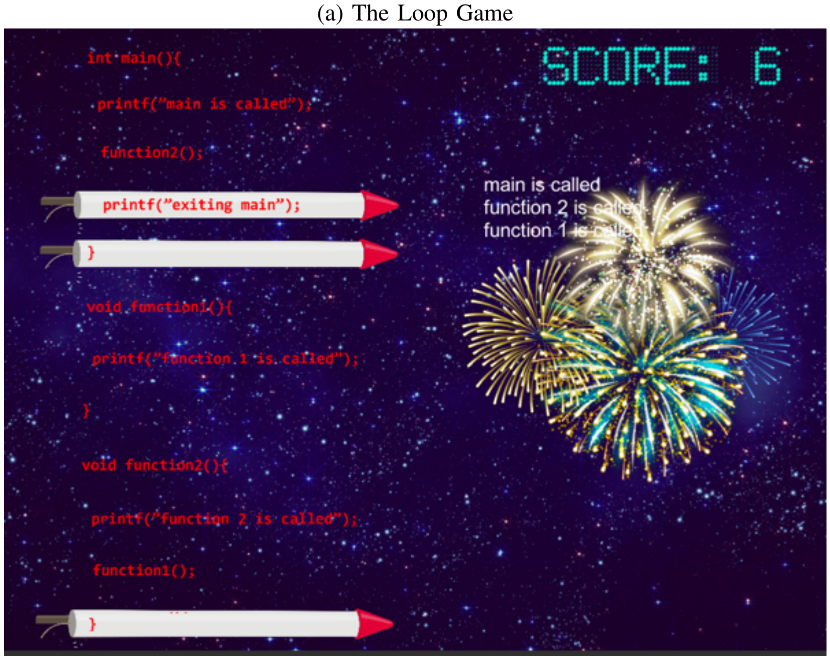
This paper explores whether additional dimensions such as learner attitude towards STEM subjects, prior educational background, and gender impact differently the attainment of learning outcomes through game-based learning.

It is widely known that women are underrepresented in the computing field. This manifests starting with formal education; for example, in the EU and US the proportion of women obtaining a Bachelor's Degree in computing courses is between 18-21% [22], [23]. To address the gender inequity, it is crucial to understand the rationale behind the various attitudes that different genders show towards computing. Main and Schimpf's study [24] reveals that women's interest in computing are influenced by a series of factors that are changing across the different stages of life and education, from pre-high school up to employment. Their study found that the main contributors to the female's lower participation in computing in pre-high school are the differences in computer access, spanning from interests to use a computer, computing skills, to frequency and patterns of use. At the high school, the main differences between male and female students come from the courses they take and knowledge about careers in computing, with a lower number of female students taking introductory computing courses. At the college level, factors such as self-efficacy, sense of belonging, stereotypes, and classroom environments affect womens interests in pursuing computing studies. Smith *et al.* investigate the motivations that influence women to study computing by conducting a student survey [25]. Their findings show that the main factors for womens participation in computing degrees are early experiences with computing, interest in computing, prospects at a good career, family members working in the computing area, and advice received from family members or school advisers. Potvin *et al.* show that women's interests in computing are correlated with their career outcome expectations [26].

Recent efforts to reduce the gender gap in computer programming propose teaching programming to primary and secondary school pupils via the design and development of serious games [27].

However, there is limited research on the role gender play in the effectiveness of using serious games to facilitate students performance and learning. Buffum *et al.* [28] combine collaborative learning and game-based learning to enable students who have less prior gaming experience to better engage with the game, and thus facilitating learning computational thinking at middle school level for underrepresented groups. In their study the authors consider underrepresentation based on gender and differences from prior experience with programming and video gaming. The findings show that while female students had lower knowledge gain than the male students in the first level of the game, the more the students interacted with the game the lower the gender learning gap. Through collaboration the students gain significant knowledge regardless of their gender or prior gaming experience. While female students may have less prior gaming experience than male students, it was not investigated if similar patterns regarding knowledge gain are observed for male students without prior gaming experience.

(a) The Loop Game



(b) The Function Game

A recent study [29] explored the factors that influence the engagement with an educational game for mathematics. While the rewards offered within the game is the most important factor for both male and female students, the other factors that attract them to engage with the game differ. The male students specified that they favor games that are challenging and provide feedback, whereas female students preferred games with a clear goal and with elements of social interaction. The study was conducted with young students (i.e. between 8-11 years), therefore the findings may not hold for different age categories. Furthermore, the study did not assess whether the effectiveness of using the game for knowledge acquisition is different depending on gender. In this context, more studies are needed to better gauge if game-based learning provides an equitable learning experience and knowledge gain for both male and female students.

This paper examines the influence of learners' attitude towards STEM subjects and their prior educational background in improving learners' knowledge and learning outcomes when using educational games. More specifically, the current study investigates whether learners with different demographics show different degrees of knowledge gain in programming when playing serious games.

III. NEWTON Project and Serious Games

A. NEWTON Project

The NEWTON project is a large EU Horizon 2020 project which designed, developed and deployed innovative solutions for technology-enhanced learning involving delivery of state-of-the-art STEM content to diverse learner audiences [30]. The NEWTON solutions include innovative technologies for adaptive and personalised multimedia, multi-sensorial media delivery, Virtual and Augmented Reality (VR/AR) learning, Virtual Teaching and Learning Labs (Virtual Labs), Fabrication Labs (Fab Labs) and Gamification-based teaching and learning. These solutions are used in conjunction with different pedagogical approaches including self-directed, game-based and problem-based learning methods. The NEWTON project has also designed and developed a new learning management platform, NEWTELP[2], which embeds the NEWTON innovative solutions. Both the NEWTON educational content and NEWTELP platform have been tested in 20 primary, secondary and third level institutions, including in schools with students with special educational needs, across 6 different EU countries. For example, the authors of [31] presented the results of a NEWTON project Fab Lab pilot which improved students' learning experience, and follow up work [3] showed that Fab Lab learning is effective in fostering students' interest in STEM. The authors of [32] discussed the results of a pilot investigating the use of VR and Virtual Labs when teaching STEM subjects in primary schools. Employing adaptive multimedia content and delivering multiple sensorial media content in educational contexts was reported in [33] and [34], respectively. Game-based learning was also performed and the positive results obtained are reported in [15]. Finally NEWTON technologies were used to deliver content to both normal development learners and learners with special educational needs, as reported in [35].

Relevant to this paper is that the serious games described here are part of the NEWTON project and were deployed on the NEWTELP platform. The evaluation procedure, which follows a methodology resulted from the research performed within the NEWTON project was also deployed in NEWTELP.

B. Programming Pilot Serious Games

In order to improve students' learning experience and learning outcome, three 2D serious games that visualize different topics in C Programming were developed as part of the NEWTON Project. These topics were identified by academics as among those most difficult for students to understand: *loops*, *functions*, and *structures*.

1) The Loop Game: In the 3-level Loop Game, students learn about the concepts of basic *for-loops*, *for-loops* with `continue` statement and *for-loops* with `break` statement through an interactive 2D undersea scenario, as illustrated in Fig. 1a. In this game, the player plays the role of a mermaid who carries out repetitive tasks of collecting and storing coins into a treasure chest. In level 1, the mermaid needs to collect and store five coins in total, which corresponds to a normal *for-loop* with 5 repetitions. In level 2, the same amount of coins need to be collected, however, some coins would disappear when collected by the mermaid. This situation illustrates the effects of a `continue` statement as the mermaid would skip the remaining routine of storing the coin into the treasure chest and continue to collect the next coin instead. In the last level, one of the 5 coins would turn into a jackpot once collected, with the whole task being finished immediately. This design effectively visualizes the effects of a `break` statement inside a *for-loop*. In all three levels, the source code is displayed at the left hand side of the game with lines that are currently being executed highlighted in red, therefore enabling code tracing.

2) The Function Game: The Function Game, as illustrated in Fig. 1b visualizes the concepts of functions, the execution order of statements and ways of passing parameters in C Programming through an interactive firework event scenario. In each level, the player is asked to fire the firework rockets according to the execution order of the statements that each rocket carries. One point is awarded for each correct action, whereas the same amount is deducted upon each wrong action. There is also a question in each level, which attracts extra points. The 3 levels cover the basic functions, *passing-parameters-by-value* and *passing-parameters-by-address*, respectively. Key knowledge concepts of each line of code is displayed at the right hand side of the screen when the corresponding firework rocket is fired.

[2]NEWTELP platform webpage, http://newtelp.eu

3) The Structure Game: In the Structure Game, as illustrated in Fig. 1c, the abstract concepts of structure in C Programming are visualized in an easy-to-understand restaurant scenario. In the first level of this 3-level drag-and-drop game, students learn the definition of *structure*, the syntax for declaring a `struct` with tag name, defining type name with `typedef` keyword through conducting the tasks of creating a pizza set menu (i.e., a pizzaSet structure) and a sushi set menu (i.e., a sushiSet structure) from individual food/drink types such as pizza, sushi, dessert, coffee, and drinks, which are treated as basic data types in C. In the second and third level of the game, students learn the syntax of declaring a *structure* variable using structure tag name or using structure type name and accessing members using member access operator through ordering his/her own sushi set and pizza set. Throughout the game, a restaurant waiter character explains all the relevant programming knowledge details as the player progresses.

IV. The Programming Pilot

A. Pilot Setup and Participants

The impact of the three proposed serious games on students' learning outcomes were evaluated during the NEWTON Project Programming large scale pilot deployed at Dublin City University-Ireland. This pilot tested various TEL materials, such as serious games, problem based learning, adaptive multimedia content, which were integrated as part of the EM108 Software Development for Engineers module. The participants of the pilot were first year undergraduate students from the Faculty of Engineering and Computing who enrolled in the module, while most of them have little or no prior programming experience. All learning materials, including the serious games and associated tests, were deployed on the NEWTELP Platform, as illustrated in Figure 2.

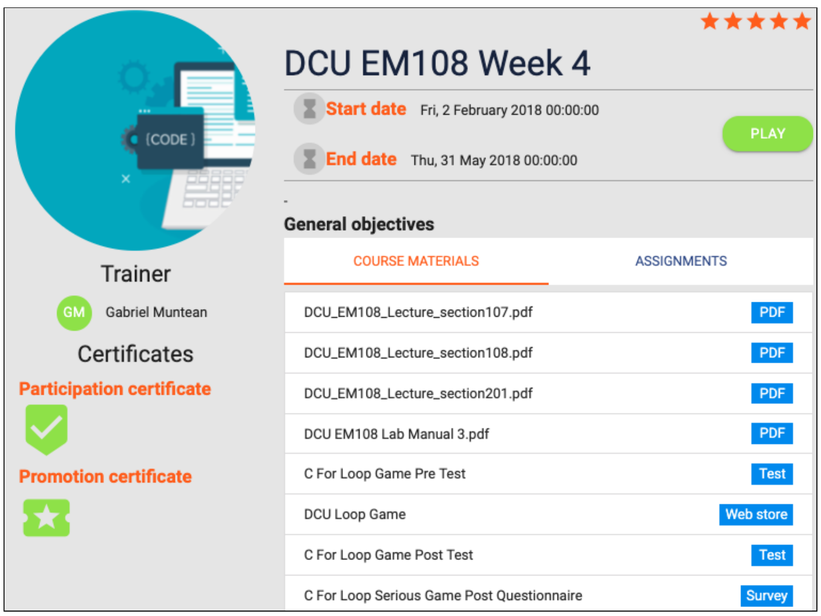


Fig. 2: Screenshot of serious game on the NEWTELP platform

B. Evaluation Methodologies

A multi-dimensional pedagogical and learner attitude assessment toolkit [36] was proposed by the NEWTON pedagogical assessment committee to comprehensively evaluate the outcomes of all NEWTON pilots. Following the toolkit, a series of questionnaires and tests were adopted throughout the pilot. At the beginning of the semester, the students were asked to complete a demographic questionnaire and a learner attitude and affective state pre-questionnaire through the NEWTELP online platform, prior to their interactions with any NEWTON technologies. In particular, the demographic questionnaire collects general information of participants in terms of demographics, their attitude towards learning and their overall use of technology. During the pilot, the serious games were integrated as part of the lab assignments of the corresponding topics. Students experienced the games in the computer laboratories during the lab sessions, which were scheduled on the second day following the lecture sessions. For each game, a pair of pre- and post-tests with the same number (i.e. 3 or 4) of single choice questions targeting the knowledge topics covered in the game, were conducted immediately prior and after their interaction with the game to evaluate students' knowledge gains by playing the game. A post-game questionnaire that surveyed students' experience of the game in terms of usability, perceived knowledge acquisition, and general user experience was also conducted after the post-test. Towards the end of the pilot, participants were asked to fill in a learner attitude and affective state post-questionnaire, which, together with the pre-questionnaire, could reveal the changes in students' interests in the subject through using NEWTON technologies.

This paper focuses on analyzing students' knowledge gain by considering both the results of demographic questionnaire, and pre- and post-tests for all games. The students are classified into four demographic subgroups based on their answers to certain questions in the demographic questionnaire:

1) *Educational ability subgroups* - determined by participants' answers to "Do you get good marks in science, technology and maths?"
   - Good marks subgroup: includes participants who answered "Yes, always" or "Yes, sometimes".
   - Average marks subgroup: includes participants who answered "Average marks".
   - Bad marks subgroup: includes participants who answered "Low marks" or "Terrible marks".
2) *Initial attitude to university subgroups* - determined by participants' answers to "How do you feel about university?"
   - Love/like university subgroup: includes participants who answered "I love it" or "I like it".
   - University is OK subgroup: includes participants who answered "It is OK".
   - Don't like university subgroup: includes participants who said "I don't like it" or "I don't like it at all".
3) *Initial attitude towards learning STEM subgroups* - determined by participants' answers to "How do you feel about learning science, technology and maths?"
   - Love/like learning STEM subgroup: includes participants who answered "I love it" or "I like it".
   - Learning STEM is OK subgroup: includes participants who answered "It is OK".
   - Don't like learning STEM subgroup: includes participants who answered "I don't like it" or "I don't like it at all".
4) *Gender subgroups* - determined by participants' answers to "Are you male or female?"
   - Male subgroup: includes male participants
   - Female subgroup: includes female participants

Out of the 133 students enrolled in the EM108 Software Development for Engineers module, 87 students filled in the Demographic Questionnaire. Therefore, only the 87 students are considered in the following analysis. As the participants were third level education students, 34% were aged under 19, 63% were aged between 19 and 24 and the remaining 3% were aged 25 or above. The methodology and procedure of the pilot have obtained ethical approval from the DCU Research Ethics Committee and have followed the Code of Ethical Practice set out by the university. All participants were presented with plain language test description and data management plan documents and had signed informed consent forms before their results have been analyzed and have been included in this paper. It is worth noting that due to students' absence, for various reasons, in different lab sessions the numbers of students (see Section V) who participated in each game varies across the three games.

V. Knowledge Assessment Result Analysis

This section presents the knowledge assessment results obtained in the pilot for the three proposed games. The results considering all participants are analyzed first and then analysis of tests results for different subgroups is conducted.

A. Overall Results

The average pre- and post-test marks among all students for each game are presented in Table I and Figure. 3. Table II outlined the paired-sample t-test of the pre- and post-test results of all students in each game. As can be seen, statistically significantly improved learning outcomes were obtained with all three games among all students.

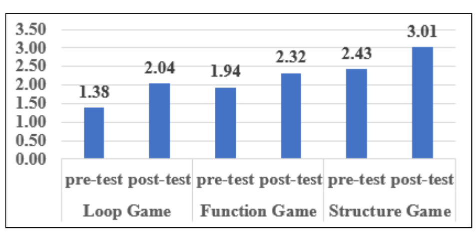


Fig. 3: Comparison of average pre- and post-test marks among all students for each game

In the Loop Game (N=55), the average post-test score (2.04 out of 3) showed an improvement of 0.655 compared with the average pre-test score (1.38 out of 3); a paired-sample t-test confirmed statistical significance of these results with $\alpha = 0.05$ ($t(54) = 4.926$, $p = 0.000$).

In the Function Game (N=68), the average post-test score (2.32 out of 3) showed an improvement of 0.382 compared with the average pre-test score (1.94 out of 3); a paired-sample t-test confirmed statistical significance of these results with $\alpha = 0.05$ ($t(67) = 3.040$, $p = 0.003$).

In the Structure Game (N=67), the average post-test score (3.01 out of 4) showed an improvement of 0.582 compared with the average pre-test score (2.43 out of 4); a paired-sample t-test confirmed statistical significance of these results with $\alpha = 0.05$ ($t(66) = 3.417$, $p = 0.001$).

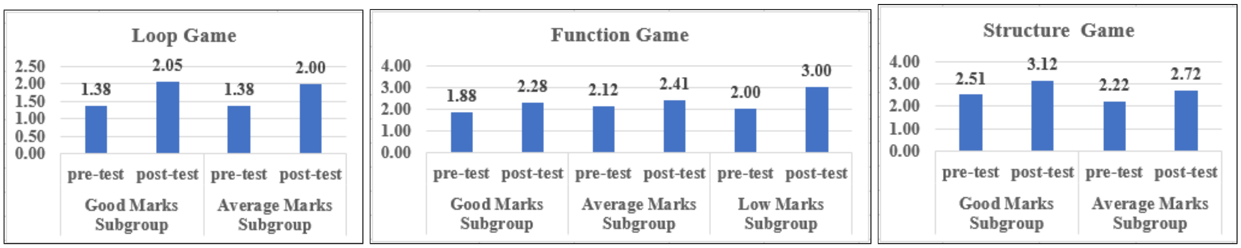


Fig. 4: Comparison of average pre- and post-test marks among different educational ability subgroups in each game

TABLE I: Average pre- and post-test marks for all students

Game	Test	Mean	N	Std Dev (SD)	Std Error Mean (SE)
Loop	pre	1.38	55	0.972	0.131
	post	2.04	55	0.816	0.110
Function	pre	1.94	68	0.862	0.105
	post	2.32	68	0.854	0.104
Structure	pre	2.43	67	1.362	0.166
	post	3.01	67	1.022	0.125

TABLE II: Paired-sample t-test of learning outcome improvements among all students (CI is short for Confidence Interval)

Game	Paired Differences					t	df	Sig. (2-tailed)
	Mean	SD	SE	95% CI of the Difference				
				Lower	Upper			
Loop	0.655	0.985	0.133	0.388	0.921	4.926	54	0.000
Function	0.382	1.037	0.126	0.131	0.633	3.040	67	0.003
Structure	0.582	1.394	0.170	0.242	0.922	3.417	66	0.001

B. Educational Ability-based Results

The average pre- and post-test marks among students belong to different educational ability subgroups for each game are presented in Table III and Figure. 4. Table IV outlined the paired-sample t-test of the pre- and post-test results of students in each educational ability subgroup in each game. Improved learning outcomes were observed in all educational subgroups in all games, while the improvements in some subgroups are statistically significant.

The Good Marks subgroup obtained improved learning outcomes in all three games, and all results are statistically significant with $\alpha = 0.05$. The Loop Game achieved statistically significant improvement ($t(41) = 4.528$, $p = 0.000$) in the post-test score (2.05 out of 3) compared with the the pre-test score (1.38 out of 3). The Function Game achieved statistically significantimprovement ($t(49) = 2.556$, $p = 0.0014$) in the post-test score (2.28 out of 3) compared with the the pre-test score (1.88 out of 3). The Structure Game achieved statistically significant improvements ($t(48) = 3.102$, $p = 0.003$) in post-test marks (3.12 out of 4) compared with pre-test scores (2.51 out of 4).

The Average Marks subgroup resulted in improved learning outcomes in all games, but none of the improvements is statistically significant with $\alpha = 0.05$.

There is only one student that claimed to have low marks. S/he only attended the Function Game session and got improved learning outcome (3 in post-test, 2 in pre-test). Since the sample size is too small, t-test was not conducted.

C. Attitude towards University-related Results

Table V and Figure 5 show the average pre- and post-tests marks of students belonging to different initial-attitude-towards-school subgroups while Table VI further exploits the paired-sample t-test results to reveal whether the difference between the average pre- and post-test marks in those subgroups are significant or not.

As can be observed from Table V, all subgroups achieved improved learning outcomes after interacting with each game.

In particular, the Love/like-university subgroup achieved statistically significant learning outcome improvements with each game ($\alpha = 0.05$). In the Loop Game, this subgroup obtained statistically significant improvement ($t(45) = 4.210$, $p = 0.000$)

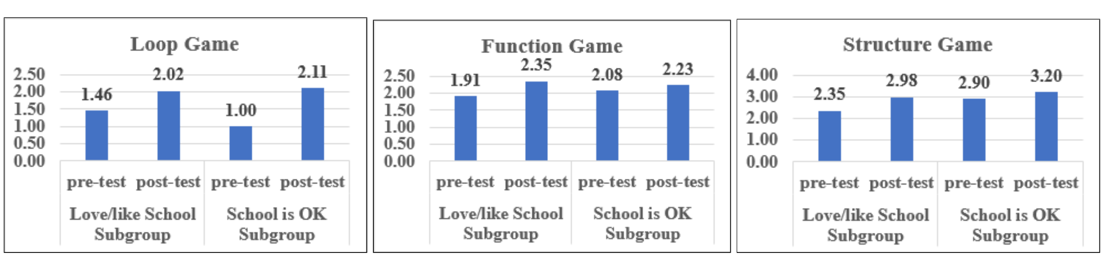


Fig. 5: Comparison of average pre- and post-test marks for students with different attitudes towards school

TABLE III: Average pre- and post-test marks in different educational-ability subgroups

Subgroup	Game	Test	Mean	N	SD	SE
Good Marks	Loop	pre	1.38	42	0.962	0.148
		post	2.05	42	0.825	0.127
	Function	pre	1.88	50	0.918	0.130
		post	2.28	50	0.882	0.125
	Structure	pre	2.51	49	1.431	0.204
		post	3.12	49	1.033	0.148
Average Mark	Loop	pre	1.38	13	1.044	0.290
		post	2.00	13	0.816	0.226
	Function	pre	2.12	17	0.697	0.169
		post	2.41	17	0.795	0.193
	Structure	pre	2.22	18	1.166	0.275
		post	2.72	18	0.958	0.226
Low Marks	Loop	pre	.	0	.	.
		post	.	0	.	.
	Function	pre	2	1	.	.
		post	3	1	.	.
	Structure	pre	.	0	.	.
		post	.	0	.	.

TABLE IV: Paired-sample t-test of learning outcome improvements in different educational-ability subgroups

Subgroup	Game	Paired Differences					t	df	Sig. (2-tailed)
		Mean	SD	SE	95% CI of the Difference				
					Lower	Upper			
Good Marks	Loop	0.667	0.954	0.147	0.369	0.964	4.528	41	0.000
	Function	0.400	1.107	0.156	0.086	0.714	2.556	49	0.014
	Structure	0.612	1.382	0.197	0.215	1.009	3.102	48	0.003
Average Marks	Loop	0.615	1.121	0.311	-0.062	1.293	1.979	12	0.071
	Function	0.294	0.849	0.206	-0.142	0.731	1.429	16	0.172
	Structure	0.500	1.465	0.345	-0.229	1.229	1.448	17	0.166

TABLE V: Average pre- and post-test marks in different initial-attitude-towards-university subgroups

Subgroup	Game	Test	Mean	N	SD	SE
Love/like university	Loop	pre	1.46	46	0.959	0.141
		post	2.02	46	0.830	0.122
	Function	pre	1.91	55	0.867	0.117
		post	2.35	55	0.844	0.114
	Structure	pre	2.35	57	1.343	0.178
		post	2.98	57	0.973	0.129
University is OK	Loop	pre	1.00	9	1.000	0.333
		post	2.11	9	0.782	0.261
	Function	pre	2.08	13	0.862	0.239
		post	2.23	13	0.927	0.257
	Structure	pre	2.90	10	1.449	0.458
		post	3.20	10	1.317	0.416

TABLE VI: Paired-sample t-test of learning outcome improvements in different initial-attitude-towards-university subgroups

Subgroup	Game	Paired Differences					t	df	Sig. (2-tailed)
		Mean	SD	SE	95% CI of the Difference				
					Lower	Upper			
Love/like university	Loop	0.565	0.910	0.134	0.295	0.836	4.210	45	0.000
	Function	0.436	1.032	0.139	0.157	0.715	3.135	54	0.003
	Structure	0.632	1.459	0.193	0.244	1.019	3.267	56	0.002
University is OK	Loop	1.111	1.269	0.423	0.135	2.087	2.626	8	0.030
	Function	0.154	1.068	0.296	-0.492	0.799	0.519	12	0.613
	Structure	0.300	0.949	0.300	-0.379	0.979	1.000	9	0.343

in the post-test score (2.02 out of 3) compared with the pre-test result (1.46 out of 3). In the Function Game, this subgroup achieved statistically significant improvement ($t(54) = 3.135$, $p = 0.003$) in the post-test score (2.35 out of 3) compared with the pre-test score (1.91 out of 3). In the Structure Game, the improvements were also statistically significant ($t(56) = 3.267$, $p = 0.002$) in the post-test (2.98 out of 4) compared with the pre-test (2.35 out of 4).

For the University-is-OK subgroup, only one game resulted in statistically significant improved learning outcome with $\alpha = 0.05$, i.e., the Loop Game ($t(8) = 2.626$, $p = 0.030$), while the other two games saw improved learning outcomes that were not statistically significant.

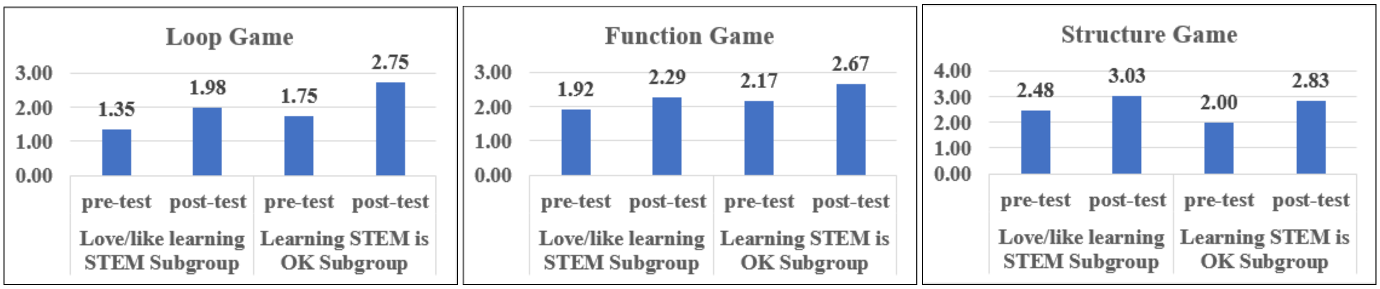


Fig. 6: Comparison of average pre- and post-test marks among students with different attitudes towards STEM learning

TABLE VII: Average pre- and post-test marks in different attitude-towards-learning-STEM subgroups

Subgroup	Game	Test	Mean	N	SD	SE
Love/like learning STEM	Loop	pre	1.35	51	0.934	0.131
		post	1.98	51	0.812	0.114
	Function	pre	1.92	62	0.836	0.106
		post	2.29	62	0.857	0.109
	Structure	pre	2.48	61	1.349	0.173
		post	3.03	61	0.983	0.126
Learning STEM is OK	Loop	pre	1.75	4	1.500	0.750
		post	2.75	4	0.500	0.250
	Function	pre	2.17	6	1.169	0.477
		post	2.67	6	0.816	0.333
	Structure	pre	2.00	6	1.549	0.632
		post	2.83	6	1.472	0.601

D. Attitude towards Learning STEM-related Results

In this subsection, the learning outcomes in each game among students who belong to different initial-attitude-towards-learning-STEM subgroups are investigated. As indicated by Table VII and Figure 6, again, both the Love/like-learning-STEM subgroup and the Learning-STEM-is-OK subgroup obtained improved learning outcomes with every game. Paired-sample t-test results, as shown in Table VIII, further examined the significance of the knowledge gains in each subgroup.

The Love/like-learning-STEM subgroup achieved statistically significant learning outcome improvements with each game with $\alpha = 0.05$. In the Loop Game, this subgroup obtained statistically significant ($t(50) = 4.676$, $p = 0.000$) improvement in the post-test score (1.98 out of 3) compared with the the pre-test score (1.35 out of 3). In the Function Game, this subgroup achieved statistically significant ($t(61) = 2.718$, $p = 0.009$) improvement in the post-test score (2.29 out of 3) compared with the the pre-test score (1.92 out of 3). In the Structure Game, this subgroup achieved statistically significant ($t(60) = 3.040$, $p = 0.004$) improvement in the post-test score (3.03 out of 4) compared with the the pre-test score (2.48 out of 4).

For the Learning-STEM-is-OK subgroup, all games resulted in knowledge gain, but not statistically significant ($\alpha = 0.05$).

TABLE VIII: Paired-sample t-tests of learning outcome improvements for different initial-attitude-to-learning-STEM

Subgroup	Game	Paired Differences					t	df	Sig. (2-tailed)
		Mean	SD	SE	95% CI of the Difference				
					Lower	Upper			
Love/like learning STEM	Loop	0.627	0.958	0.134	0.358	0.897	4.676	50	0.000
	Function	0.371	1.075	0.137	0.098	0.644	2.718	61	0.009
	Structure	0.557	1.432	0.183	0.191	0.924	3.040	60	0.004
Learning STEM is OK	Loop	1.000	1.414	0.707	-1.250	3.250	1.414	3	0.252
	Function	0.500	0.548	0.224	-0.075	1.075	2.236	5	0.076
	Structure	0.833	0.983	0.401	-0.198	1.865	2.076	5	0.093

E. Gender-related Results

In this subsection, the impacts of gender on students' learning outcomes with games are examined. As can be observed from Table IX and Figure 7, knowledge gains were obtained in both male and female subgroups with every game.

Results from Table X show that the knowledge gains in two out the three games, i.e. Loop and Function games, for the male subgroup were statistically significant. In the Loop Game, male students obtained statistically significant improvements ($t(45) = 4.451$, $p = 0.000$) in the post-test score (2.04 out of 3) compared with the pre-test (1.39 out of 3). In the Function Game, male students achieved better statistically significant post-test results ($t(53) = 3.522$, $p = 0.001$) (2.43 out of 3) compared with those in pre-test (1.94 out of 3). In the Structure Game, however, the knowledge gain from pre- to post-test among male students was not statistically significant with $\alpha = 0.05$.

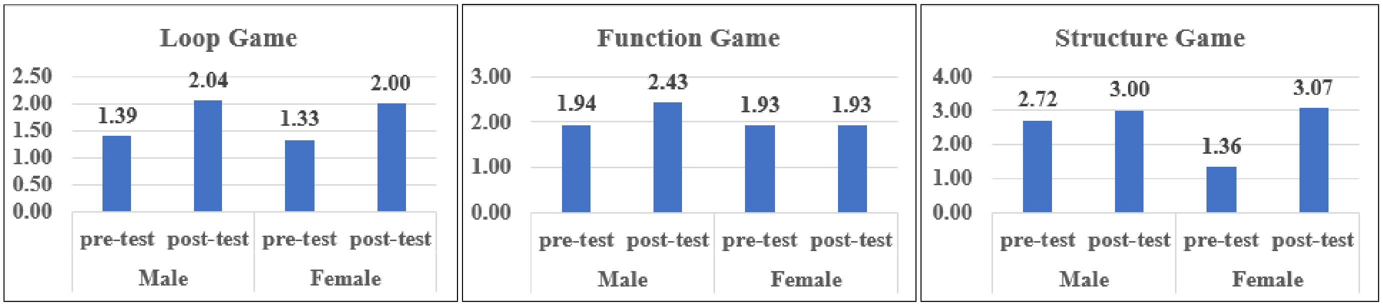


Fig. 7: Comparison of average pre- and post-test marks for male and female students for each game

TABLE IX: Average pre- and post-test marks in different gender subgroups

Subgroup	Game	Test	Mean	N	SD	SE
Male	Loop	pre	1.39	46	1.043	0.154
		post	2.04	46	0.842	0.124
	Function	pre	1.94	54	0.878	0.119
		post	2.43	54	0.792	0.108
	Structure	pre	2.72	53	1.321	0.181
		post	3.00	53	1.056	0.145
Female	Loop	pre	1.33	9	0.500	0.167
		post	2.00	9	0.707	0.236
	Function	pre	1.93	14	0.829	0.221
		post	1.93	14	0.997	0.267
	Structure	pre	1.36	14	0.929	0.248
		post	3.07	14	0.917	0.245

TABLE X: Paired-sample t-test of learning outcome improvements in different gender subgroups

Subgroup	Game	Paired Differences					t	df	Sig. (2-tailed)
		Mean	SD	SE	95% CI of the Difference				
					Lower	Upper			
Male	Loop	0.652	0.994	0.147	0.357	0.947	4.451	45	0.000
	Function	0.481	1.005	0.137	0.207	0.756	3.522	53	0.001
	Structure	0.283	1.306	0.179	-0.077	0.643	1.577	52	0.121
Female	Loop	0.667	1.000	0.333	-0.102	1.435	2.000	8	0.081
	Function	0.000	1.109	0.296	-0.641	0.641	0.000	13	1.000
	Structure	1.714	1.139	0.304	1.057	2.372	5.633	13	0.000

Female students achieved statistically significant knowledge gains in one game, i.e., the Structure Game ($\alpha = 0.05$). In this game, the improvement was from an average mark of 1.36 out of 4 in the pre-test to 3.07 out of 4 in the post-test ($t(13) = 5.633$, $p = 0.000$). Female students achieved improved learning outcomes with the Loop Game as well, but they were not statistically significant with $\alpha = 0.05$. In the Function Game, however, the average post-test results were similar with those of the pre-test.

VI. Knowledge Assessment Discussions

Section V investigated the knowledge gains through each game among all students as well as among different subgroups identified by angles such as educational ability, initial attitude to university, initial attitude towards learning STEM subjects, and gender. As the results reveal, the piloted games were able to facilitate students' understanding of the target Programming knowledge and improve their performance in tests. Students with different attitude and affective states do show different patterns in knowledge gain through games. Students with a stronger attitude towards STEM and affective state level, i.e., those who get good marks in STEM subjects, who love/like school and who love/like learning STEM subjects, tend to gain more through games, as they achieved statistically significant improvements in the knowledge tests in every game. Students who had an average attitude and affective state level, i.e., those who get average marks in STEM subjects, who thought school is OK and who believe learning STEM subjects is OK, still got improved learning outcomes in all games, although the improvements were not always statistically significant. Such phenomenon is understandable, considering the fact that students with stronger attitude towards STEM subjects and higher affective state level tend to be better at learning regardless of learning manner.

VII. Conclusion

This paper presented three serious games introduced to help students improve their understand of fundamental C Programming concepts. The impact of these games on students' learning outcome was studied in a large scale pilot carried out as part of a 12 week university software engineering module. Learning outcomes were assessed considering all students and students grouped based on their educational ability, attitude to university or learning STEM, and gender. The results show that in general students have achieved statistically significant knowledge gains. Furthermore, students with better motivation and affective state tend to obtain more significant knowledge gain via serious games, compared with other students. Future work will use these results as input to personalization in order to tailor learning to students' personal preferences and improve both learning experience and learning outcome.

Acknowledgment

This work was supported by the EU Horizon 2020 Research and Innovation programme via grant no. 688503 for NEWTON (http://NEWTONproject.eu). The Science Foundation Ireland's support via grant 13/RC/2094 to Lero - the SFI Research Centre for Software - is also gratefully acknowledged.

References

[1] M. Zyda, "From visual simulation to virtual reality to games," *Computer*, vol. 38, pp. 25–32, Sept. 2005.

[2] D. Djaouti, J. Alvarez, and J.-P. Jessel, "Classifying serious games: the g/p/s model," in *Handbook of research on improving learning and motivation through educational games: Multidisciplinary approaches*, pp. 118–136, IGI Global, 2011.

[3] M. A. Togou, C. Lorenzo, G. Cornetta, and G.-M. Muntean, "Assessing the effectiveness of using fab lab-based learning in schools on k12 students attitude toward steam," *IEEE Trans. Edu.*, vol. 63, no. 1, 2020.

[4] P. Backlund and M. Hendrix, "Educational games – are they worth the effort? a literature survey of the effectiveness of serious games," in *Proc. 5th Int. Conf. Games Virtual Worlds Serious Applicat. (VS-GAMES)*, pp. 1–8, Sept. 2013.

[5] E. A. Boyle, T. Hainey, T. M. Connolly, G. Gray, J. Earp, M. Ott, T. Lim, M. Ninaus, C. Ribeiro, and J. Pereira, "An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games," *Comput. & Edu.*, vol. 94, pp. 178–192, 2016.

[6] J. Hamari, D. J. Shernoff, E. Rowe, B. Coller, J. Asbell-Clarke, and T. Edwards, "Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning," *Comput. Human Behavior*, vol. 54, pp. 170–179, 2016.

[7] C. H. Muntean, J. Andrews, and G.-M. Muntean, "Final frontier: an educational game on solar system concepts acquisition for primary schools," in *Proc. IEEE 17th Int. Conf. Adv. Learning Technol. (ICALT)*, pp. 335–337, Jul. 2017.

[8] C. H. Muntean, N. E. Mawas, M. Bradford, and P. Pathak, "Investigating the impact of an immersive computer-based math game on the learning process of undergraduate students," in *Proc. IEEE Frontiers Educ. Conf. (FIE)*, pp. 1–8, Oct. 2018.

[9] C. C. Liu, Y. B. Cheng, and C. W. Huang, "The effect of simulation games on the learning of computational problem solving," *Comput. & Educ.*, vol. 57, pp. 1907–1918, Nov. 2011.

[10] I. Ghergulescu and C. H. Muntean, *Measurement and Analysis of Learner's Motivation in Game-Based E-Learning*, pp. 355–378. New York, NY, USA: Springer New York, 2012.

[11] M. Eagle and T. Barnes, "Experimental evaluation of an educational game for improved learning in introductory computing," in *Proc. 40th ACM Tech. Symp. Comput. Sci. Educ. (SIGCSE'09)*, pp. 321–325, Mar. 2009.

[12] N. Adamo-Villani, T. Haley-Hermiz, and R. Cutler, "Using a serious game approach to teach 'operator precedence' to introductory programming students," in *Proc. 17th Int. Conf. on Inform. Vis.*, pp. 523–526, 2013.

[13] J. Tessler, B. Beth, and C. Lin, "Using cargo-bot to provide contextualized learning of recursion," in *Proc. 9th Annual Int. ACM Comput. Educ. Research Conf. (ICER)*, pp. 161–168, 2013.

[14] A. Mathrani, S. Christian, and A. Ponder-Sutton, "Playit: game based learning approach for teaching programming concepts," *J. Educational Technol. & Soc.*, vol. 19, no. 2, pp. 5–17, 2016.

[15] D. Zhao, A. E. Chis, G.-M. Muntean, and C. H. Muntean, "A large-scale pilot study on game-based learning and blended learning methodologies in undergraduate programming courses," in *Proc. 10th Int. Conf. Educ. New Learning Technol. (EDULEARN)*, pp. 3716–3724, Jul. 2018.

[16] D. Zhao and G.-M. Muntean, "The newton project warehouse game: A variable and data type serious game for c programming courses," in *Proc. 10th Int. Conf. Educ. New Learning Technol. (EDULEARN)*, pp. 3709–3714, Jul. 2018.

[17] M. A. Miljanovic and J. S. Bradbury, "A review of serious games for programming," in *Serious Games*, pp. 204–216, Springer International Publishing, 2018.

[18] C. Kazimoglu, M. Kiernan, L. Bacon, and L. Mackinnon, "A serious game for developing computational thinking and learning introductory computer programming," *Procedia-Social and Behavioral Sci.*, vol. 47, pp. 1991–1999, 2012.

[19] M. A. Miljanovic and J. S. Bradbury, "Robobug: A serious game for learning debugging techniques," in *Proc. 13th Annual ACM Int. Comput. Educ. Research Conf. (ICER)*, pp. 93–100, Aug. 2017.

[20] D. Dicheva and A. Hodge, "Active learning through game play in a data structures course," in *Proc. 49th ACM Tech. Symp. Comput. Sci. Educ. (SIGCSE)*, pp. 834–839, ACM, Feb. 2018.

[21] H. Jiau, J. Chen, and K. Ssu, "Enhancing self-motivation in learning programming using game-based simulation and metrics," *IEEE Trans. Edu.*, vol. 52, pp. 555–562, Nov. 2009.

[22] Eurostat, "Graduates by Education Level, Programme Orientation, Sex and Field of Education." https://ec.europa.eu/eurostat/web/products-datasets/product?code=EDUC_UOE_GRAD02. (accessed: 08.12.2020).

[23] W. DuBow and J. J. Gonzalez, "NCWIT Scorecard: The Status of Women in Technology." https://www.ncwit.org/sites/default/files/resources/ncwit_scorecard_data_highlights_10082020.pdf, 2020. (accessed: 08.12.2020).

[24] J. B. Main and C. Schimpf, "The underrepresentation of women in computing fields: A synthesis of literature using a life course perspective," *IEEE Trans. Edu.*, vol. 60, no. 4, pp. 296–304, 2017.

[25] S. Smith, E. Sobolewska, J. Bhardwaj, and K. Fabian, "Exploring womens motivations to study computer science," in *Proc. IEEE Frontiers Educ. Conf. (FIE)*, pp. 1–7, Oct. 2018.

[26] G. Potvin, C. McGough, L. Benson, H. J. Boone, J. Doyle, A. Godwin, A. Kirn, B. Ma, J. Rohde, M. Ross, and D. Verdn, "Gendered interests in electrical, computer, and biomedical engineering: Intersections with career outcome expectations," *IEEE Trans. Edu.*, vol. 61, no. 4, pp. 298–304, 2018.

[27] C. V. de Carvalho, . Cerar, J. Rugelj, H. Tsalapatas, and O. Heidmann, "Addressing the gender gap in computer programming through the design and development of serious games," *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, vol. 15, no. 3, pp. 242–251, 2020.

[28] P. S. Buffum, M. Frankosky, K. E. Boyer, E. Wiebe, B. Mott, and J. Lester, "Leveraging collaboration to improve gender equity in a game-based learning environment for middle school computer science," in *Proc. 2015 Research Equity Sustained Participation Eng., Comput., Technol. (RESPECT)*, pp. 1–8, Aug. 2015.

[29] O. Dele-Ajayi, R. Strachan, A. Pickard, and J. Sanderson, "Designing for all: Exploring gender diversity and engagement with digital educational games by young people.," in *Proc. IEEE Frontiers Educ. Conf. (FIE)*, pp. 1–9, 2018.

[30] C. H. Muntean, D. Bogusevschi, and G.-M. Muntean, *Innovative Technology-based Solutions for Primary, Secondary and Tertiary STEM Education*. Paragon Publishing, Sept. 2019.
[31] M. A. Togou, C. Lorenzo, E. Lorenzo, G. Cornetta, and G.-M. Muntean, "Raising students' interest in stem education via remote digital fabrication: An irish primary school case study," in *10th Int. Conf. Educ. New Learning Technol. (EDULEARN)*, pp. 2835–2840, Jul. 2018.
[32] D. Bogusevschi, C. H. Muntean, N. E. Gorji, and G.-M. Muntean, "Earth course: A primary school large-scale pilot on stem education," in *10th Int. Conf. Educ. New Learning Technol. (EDULEARN)*, Jul. 2018.
[33] T. Bi, A. A. Simiscuka, F. Rodrigues da Silva, M. A. Togou, G. Rozinaj, and G.-M. Muntean, "A study of learning experience with a dash-based multimedia delivery system," in *10th Int. Conf. Educ. New Learning Technol. (EDULEARN)*, pp. 8590–8595, Jul. 2018.
[34] T. Bi, F. Silva, G. Ghinea, and G.-M. Muntean, "Improving learning experience by employing dash-based mulsemedia delivery," in *IEEE Games, Entertainment, Media Conf. (GEM)*, pp. 1–9, 2018.
[35] I. Ghergulescu, T. Lynch, M. Bratu, A. Moldovan, C. H. Muntean, and G.-M. Muntean, "Stem education with atomic structure virtual lab for learners with special education needs," in *10th Int. Conf. Educ. New Learning Technol. (EDULEARN)*, pp. 8747–8752, Jul. 2018.
[36] L. Montandon, J. Playfoot, I. Ghergulescu, M. Bratu, D. Bogusevschi, R. Rybarova, and N. E. Mawas, "Multi-dimensional approach for the pedagogical assessment in STEM technology enhanced learning," in *Proc. EdMedia Conf.*, Jun. 2018.

Dan Zhao is a Postdoctoral Researcher with National College of Ireland. Dr Zhao obtained a BEng degree from Beijing University of Posts and Telecommunications in 2011 and a PhD degree from Queen Mary University of London in 2015. She worked for the Horizon 2020 NEWTON project since 2016. Her current research interests include innovative solutions for technology-enhanced learning.

Cristina Hava Muntean is a Senior Lecturer with School of Computing, National College of Ireland (NCI). She holds a BEng degree in Computer Science and a PhD degree awarded by Dublin City University, Ireland in 2005. Dr Cristina Muntean's research interests include adaptive and personalised technology enhanced learning, game-based learning and self-directed learning. She was the NCI coordinator of the EU Horizon 2020 project NEWTON.

Adriana E. Chis is a Lecturer with School of Computing at National College of Ireland. She received a PhD in Computer Science from University College Dublin, Ireland in 2013, and her Diploma-Engineer (Honors) title in Computer Science and Engineering from "Politehnica" University of Timisoara, Romania in 2007. Her research interests include program analysis, compilers, runtime systems, cloud computing, and computer science education.

Gabriel-Miro Muntean is Professor with the School of Electronic Engineering at Dublin City University, Ireland, university which awarded him a PhD degree in 2004. Prof. Muntean's research interests include quality and performance-related issues of adaptive multimedia streaming, performance of wired and wireless networks and energy-aware networking.He has published over 400 papers in top-level international journals and conferences, authored four books and 22 book chapters, and edited seven additional books. His research interests include quality, performance, and energy saving issues related to multimedia and multiple sensorial media delivery, technology-enhanced learning, and other data communications over heterogeneous networks. He is an Associate Editor of the IEEE Transactions on Broadcasting and Multimedia Communications Area Editor for the IEEE Communications Surveys and Tutorials. He coordinated the EU Horizon 2020 NEWTON Project.