Psicothema 2020, Vol. 32, No. 4, 516-524
doi: 10.7334/psicothema2020.30

ISSN 0214 - 9915 CODEN PSOTEG

Effectiveness of Self-Regulation and Serious Games for Learning STEM Knowledge in Primary Education

María Consuelo Sáiz Manzanares, Sandra Rodríguez Arribas, Carlos Pardo Aguilar, and Miguel Ángel Queiruga-Dios
Universidad de Burgos

Abstract

Background: The learning of scientific and technological subjects is fundamental in the society of the 21st century. However, a gender gap is detected in the choice of degrees in these subjects. Recent studies indicate the need to take action from the primary education stage to increase student motivation towards these disciplines. **Methods:** We worked with a sample of 147 students in the final years of Primary Education. SRL and serious games were applied in initial tasks to computer programming. The objectives were to study the influence of gender, environment and academic level variables on the results in the resolution of initial programming tasks and on student satisfaction with their completion. **Results:** The mean level of results in these tasks was high (8 out of 10). However, significant differences were found for gender, academic level, and the covariate age. With respect to satisfaction, no significant differences were found except in the continuity of work. **Conclusions:** The use of SRL and serious play tasks promotes good levels of performance and satisfaction in all students, although differences in favour of the male gender are detected.

Keywords: Serious games, self-regulated learning, STEM, gender, satisfaction, computer programming.

Resumen

Efectividad de la Autorregulación y los Serious Games Para el Aprendizaje de Conocimientos STEM en Educación Primaria. **Antecedentes:** el aprendizaje de materias científico-tecnológicas es fundamental en la sociedad del s. XXI. Si bien, se detecta una brecha de género en la elección de titulaciones en estas materias. Estudios recientes indican la necesidad de realizar acciones desde la etapa de Educación Primaria para aumentar la motivación de los estudiantes hacia estas disciplinas. **Método:** se trabajó con una muestra de 147 estudiantes de los últimos cursos de Educación Primaria. Se aplicó SRL y serious games en tareas de inicio a la programación informática. Los objetivos fueron estudiar la influencia de las variables género, entorno y nivel académico sobre los resultados en la resolución de tareas de inicio a la programación y en la satisfacción de los estudiantes con su realización. **Resultados:** el nivel medio de resultados en estas tareas fue alto (8 sobre 10). Sin embargo, se hallaron diferencias significativas respecto de las variables género, nivel académico y efectos de la covariable edad. Relativo de la satisfacción no se hallaron diferencias significativas salvo en la continuidad de trabajo. **Conclusiones:** la utilización de SRL y de tareas de juego serios potencia buenos niveles de rendimiento y de satisfacción en todos los estudiantes, aunque se detectan diferencias a favor del género masculino.

Palabras clave: serious games, self-regulated learning, STEM, género, satisfacción, lenguaje de programación.

In 21st century society, education authorities have shown great interest in the enhancement of STEM (Science, Technology, Engineering and Mathematics) subjects (Stehle & Peters-Burton, 2019; Queiruga-Dios et al., 2020). The UNESCO report (Bokova, 2018) proposed a model citizen who, by 2030, will have transformative, innovative and creative thinking and skills within a digital society in situations of gender equality. This report revealed that in the 10-11 age group, boys and girls were equally committed to STEM subjects (75% and 72%, respectively). However, by the age of 18 this proportion fell to 33% for boys and 19% for girls. There is therefore a decline in the interest shown in the female gender. Subsequently, in Higher Education, the gap in the percentage of enrolment in engineering, technology, construction and computer science degrees increased for men and decreased for women (72% vs. 28%). However, the proportion tends to be equal (55% women vs. 45% men) in mathematics and statistical science degrees. It is therefore necessary to investigate the causes for this choice and its relation to the gender gap in the educational stage prior to Secondary Education.

Recent studies have shown that the gender gap already begins in Primary Education (García-Holgado, Verdugo-Castro et al., 2019). These authors recommend using forms of teaching based on experimentation. The aim is to increase motivation and improve execution, in order to try to equalize the percentages of choice between boys and girls. Regarding this objective, it has been found that the use of tasks that include gamification increases student motivation towards participation (Troiano et al., 2019). Gamification is a new concept that involves the use of what has been called serious games. These are games in which actions or

Received: January 25, 2020 • Accepted: July 13, 2020
Corresponding author: María Consuelo Sáiz Manzanares
Facultad de Ciencias de la Salud
Universidad de Burgos
09001 Burgos (Spain)
e-mail: mcsmanzanares@ubu.es

activities of reflection on a task are included. One of the reasons for their effectiveness is that they facilitate the development of metacognitive strategies of orientation and planning during the resolution of tasks (Cloude et al., 2019). These tasks have a very structured design based on a sequential order of difficulty. In addition, they include the guidance of an avatar that guides the user's learning through verbal aids and/or examples that promote Self-regulated learning (SRL) (Cerezo et al., 2010; Taub et al., 2014). This type of methodology is inserted within Advanced Learning Technologies (ALTs). Techniques that include the analysis of the user's learning process, since they record and monitor the development of the resolution within a learning loop that increases learner motivation (Zimmerman & Moylan, 2009). The objective is to promote interactive learning through a simulation structure in which aids are introduced that may be verbal, visual, or both (Patti, 2019). They must however include a precise narrative design of the activity accompanied by good technical skills, in order for serious games to produce an increase in the learners' creativity and motivation (Kretschmer & Terharen, 2019). Similarly, the use of serious games in natural teaching contexts has been shown to increase student motivation towards STEM-type tasks (Alsawaier, 2018; Bovermann & Bastiaens, 2020; Dreimane, 2019). Therefore, recent research has concluded that its use is an excellent resource in the educational context (Nogueira-Frazão & Martínez-Solana, 2019). In this way, Gallego-Durán et al. (2014) carried out a meta-analysis study that analyzed traditional teaching techniques vs. gamma teaching. These authors found that variables such as age, task type, reaction time, attention levels and motivation, among other variables, influenced learning processes. Likewise, these authors found that gamma teaching explained 72% of the variance in learning outcomes, with 45% of those outcomes specifically explained by gamma maze and puzzle tasks.

In summary, the use of serious game tasks (Clark, 1970) encourages the learner to experiment, access feedback, SRL and reduces the fear of error by increasing motivation to learn (Contreras et al., 2019). Hence, the implementation of serious games in teaching offers a current and efficient alternative to traditional teaching-learning methods. In Figure 1, an example of the work with serious games applied to learning at the beginning of programming is presented. A resource known as Blockly Games is used for this serious game, because it is an easy to use the tool on the web with free access. In addition, Blockly Games includes the figure of an avatar that guides the resolution of the tasks and that gives feedback to the student on the errors and successes during the execution.

The use of SRL also facilitates process-oriented feedback (Brooks et al., 2019; Coertjens, 2018; Hattie & Clarke, 2018; Park et al., 2019). The use of such feedback provides the learner with the opportunity to understand what is being done and what it is being done for, which enhances their autonomy in learning (Lodge et al., 2018; Sáiz-Manzanares et al., 2019). The use of this methodology in the teaching of STEM subjects has been shown to be very effective in achieving effective learning and increasing motivation (Taub et al., 2018; Zheng et al., 2020). Recent studies (Díaz-Lauzurica & Moreno-Salinas, 2019; Kintsakis & Rangoussi, 2019) have shown that the inclusion of gamma tasks in STEM subjects, specifically those related to the learning of programming skills, increases the motivation of students to perform them and improves their results. In addition, the motivation is extended to the teacher, which promotes a very effective loop between both agents (teacher-student) (Cerezo et al., 2018). The most widely used resources to teach the initial stages of programming use Scratch [Scratch is a visual programming language developed by the Lifelong Kindergarten Group of the MIT Media Lab (Marji, 2014). Its main feature is to enhance the learning of computer programming without having an in-depth knowledge of the code. Its features linked to the easy understanding of computer thinking have made it very widespread in the education of children, adolescents and adults)] and Blockly Games (Blockly Games contains a series of educational games that teach programming. It is designed for children who do not have previous experience in computer programming tasks).

In summary, the design of STEM tasks specifically in a programming language which includes SRL and serious games is of great interest to the researcher (Tucker-Raymond et al., 2019), because it can be used to visualize the application of various and especially metacognitive learning strategies during the process of solving these tasks (Sáiz-Manzanares & Marticorena-Sánchez,

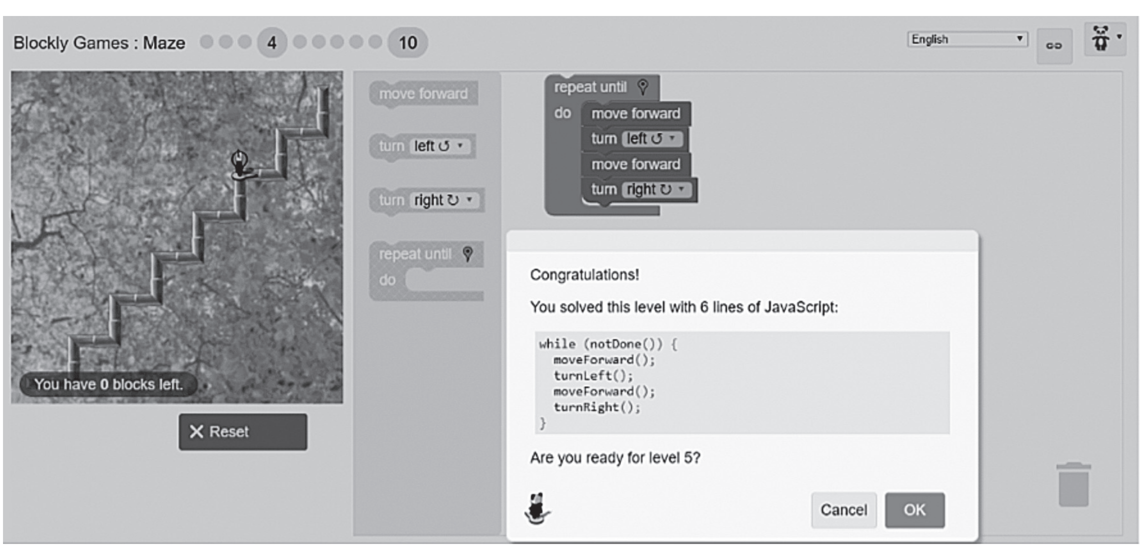


Figure 1. Example of serious games using Blockly Games

2016; Stehle & Peters-Burton, 2019; Zainal et al., 2018). In relation to this last aspect, research is currently underway to facilitate the study of the implementation of STEM experiences. These studies include the RoboSTEM project (García-Holgado, Camacho-Díaz et al., 2019) and the W-STEM project (Conde-González et al., 2019; García-Peñalvo et al., 2019). Both projects stress the need to apply the teaching of STEM subjects in natural contexts at all levels of the education system.

In view of the above, the objectives of this study were: 1) to analyse whether there were significant differences in the results obtained in the resolution of initial programming tasks with respect to the independent variables of gender, learning environment and academic level, and the covariate of age; 2) to analyse whether there were significant differences in student satisfaction with the execution of initial programming tasks in relation to the independent variables of gender, learning environment and academic level, and the covariate of age.

To answer these objectives, the following research questions were posed:

RQ1. There will be significant differences in the level of execution in tasks of initiation to computer programming carried out with serious games depending on the independent variables gender (boys vs. girls), learning environment (regular classroom vs. computer classroom) and academic year (5th vs. 6th), and the covariate age.

RQ2. There will be significant differences in student satisfaction with the execution of initial tasks to computer programming performed with serious games depending on the independent variables gender (boys vs. girls), learning environment (regular classroom vs. computer classroom), academic year (5th vs. 6th), and the covariate age.

Method

Participants

We worked with a sample of 147 students (76 boys and 71 girls) in Primary Education at the 5th and 6th grade academic levels. The characteristics of the sample can be seen in Table 1. Convenience sampling was used for the selection of the sample from educational centres whose students shared similar socioeconomic characteristics and belonged to the same city in the north of Spain.

Instruments

Blockly Games maze task. Blockly Games is a free web application that includes a series of educational games aimed at teaching programming. It is designed for children with no previous experience neither of the use of code, nor in computer programming tasks. When the execution of these games is finished, the players are ready to use the conventional text-based programming languages.

Table 1
Descriptive statistics of the participants

Academic level	N	n	Center 1 Children M	Center 1 Children SD	n	Center 1 Girls M (d)	Center 1 Girls SD (d)	n	Center 2 Children M	Center 2 Children SD (d)	Center 2 Girls M	Center 2 Girls SD (d)	n
5º	85	20	10.05	.39	24	10.17	.38	22	10.31	.57	10.21	.63	19
6º	62	24	11.04	.46	20	11.00	.45	10	11.30	.67	11	.00	8
Total	147	44			44			32					27

Note: *N* = total number of students; *n* = partial number of students; *M* = Mean average age; *SD* = age standard deviation

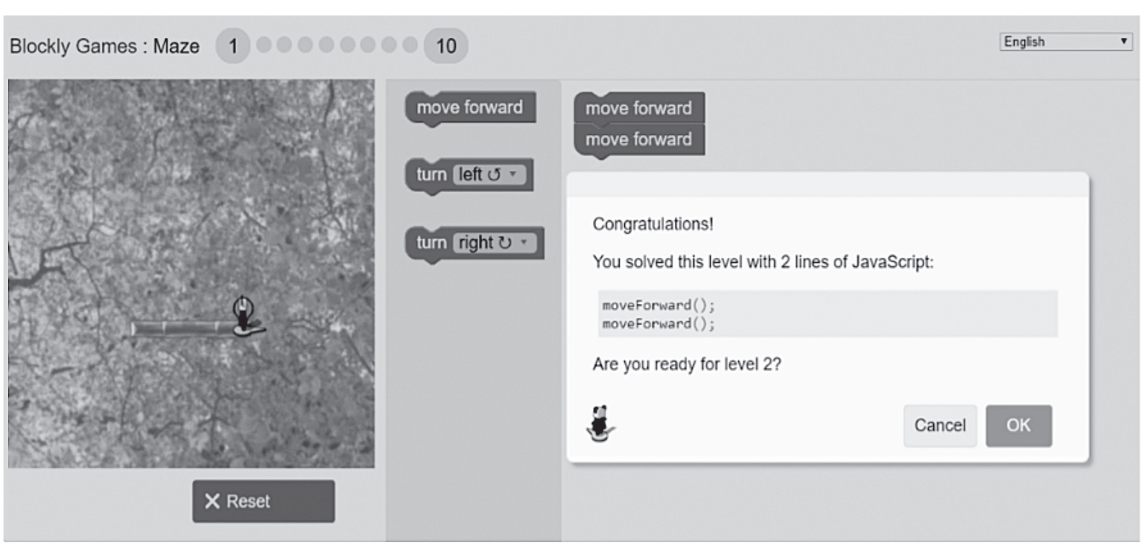


Figure 2. Feedback after completing Blockly Games maze task on level 1

Specifically, the maze task consists of solving different mazes that have degrees of increasing difficulty from 1 to 10. In such tasks, the participant can choose the type of avatar with which he will play the game (a panda bear, an astronaut or an impersonal figure). Also, before executing the activity, the trainee has to think about how it may be solved. After each game the participant receives feedback on the outcome of the execution. If the answer is correct, the programming syntax that the participant has implemented will be returned. An example of how this functions can be seen in Figure 2.

Self-Regulating instruction. In all intervention groups, Self-Regulated instruction was applied, which consisted of asking students task-oriented questions. Both declarative (related to the conceptual content of the task) and procedural (related to the planning of the resolution) metacognitive skills are used in this sort of such instruction. An example of the thinking aloud dialogue with SRL can be seen in Table 2.

Student satisfaction survey with the Blockly Games maze workshop. This instrument is an adaptation of the survey developed by Appianing and Van Eck (2018) (see Table 3), the instrument has a reliability index of α = .90. The adapted survey consists of 5 items measured on a Likert-type scale from 1 (total disagreement) to 5 (total agreement). The first question refers to whether the student has carried out similar activities and is considered as an analysis of previous knowledge.

Procedure

Prior to the start of the research, the authorization of the Bioethics Committee of the University of Burgos was obtained (IR 19/2018), as well as the authorization of the heads of the Department of Education of the Junta de Castilla y León (Spain) and the educational centres where the research was carried out. Similarly, written informed commitment was obtained from the parents or legal guardians of all the students participating in the study. Afterwards, the Blockly Games labyrinth task was worked on in the 5th and 6th grades of Primary Education. Labyrinths consists of 10 games in which students have to solve different mazes graduated in order of difficulty. At the beginning of the game, the figure of an avatar is chosen that will accompany the student during the resolution, the duration of each intervention was of one hour.

The distribution of teachers at each of the centres was respected, so the work was carried out in two environments. Environment 1: computer classroom with PC (number of students from 10 to 12) and Environment 2: ordinary classroom with Mini-PC (number of students from 21 to 25). In both environments, SRL methodology was applied. The instruction was carried out by two teachers from outside the centre, one specialised in Computer Science and the other in Psychology of Instruction.

Data analysis

A 2×2×2 factorial design was used [gender (male vs. female), learning environment (regular classroom vs. computer classroom) and academic year (5th vs. 6th)]. The dependent variables were the results in the serious games tasks (Blockly Games) and student satisfaction with them. First, a normality analysis of the sample was performed (for which asymmetry and kurtosis indicators were used), followed by an ANOVA to check whether the students from the centres were equal in the previous knowledge variable. Finally, a three-factor ANCOVA and a covariate were applied, in order to contrast the research questions, and an analysis of the value of the effect was also performed, using eta-squared (η^2). All the analyses were performed with the SPSS v.24 statistical package and a 95% confidence interval was applied.

Results

Firstly, we studied the homogeneity between the students of the two participating centres with respect to the background knowledge variable. No significant differences were found between the students of the two centres, $F(1, 147) = .00$, $p = .99$ ($\eta^2 = .00$). We then studied whether the sample met the standards of normality. To do so, the values of asymmetry and kurtosis with respect to previous knowledge were found. As can be seen in Table 4, no deviations were found in the normal values [acceptable asymmetry values should be no higher than |2.00| and those of kurtosis should not be within the interval between |8|-|20| (Bandalos & Finney, 2001)]. Based on the results, we used parametric statistics to test the research questions.

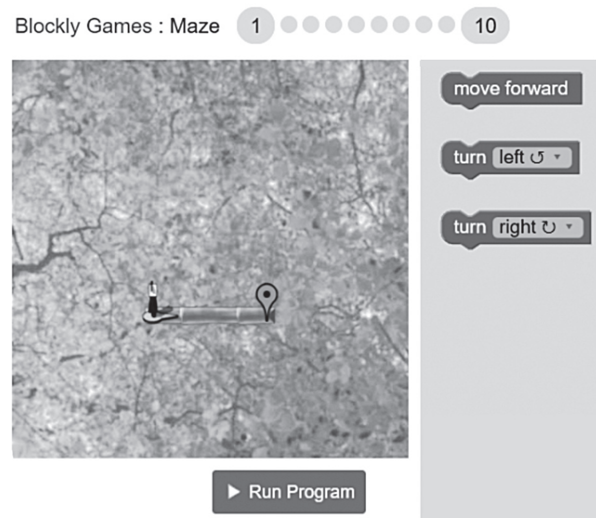
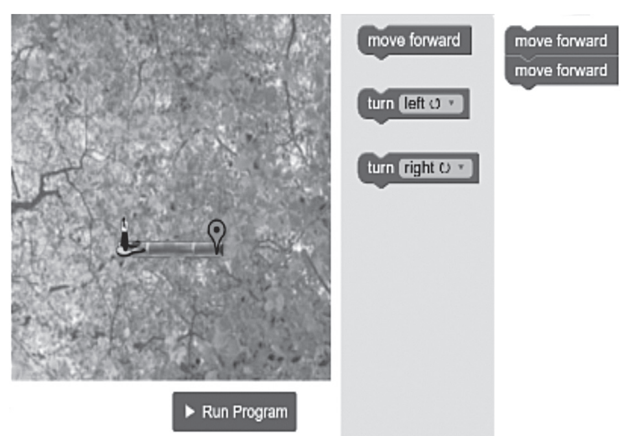
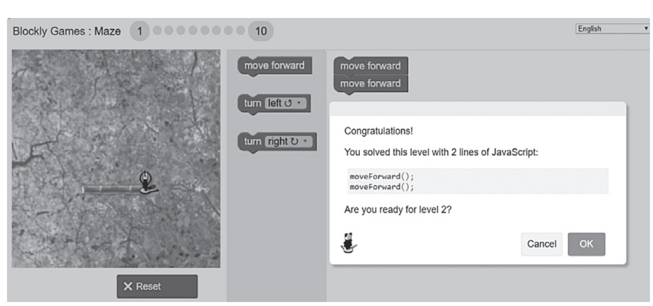
An ANCOVA with three fixed-effect factors (gender, learning environment, and academic level) was used to test RQ1, with age considered to be a covariate. Significant differences were found in the results of the execution of the Blockly Games tasks with respect to the gender variable $F(1, 147) = 4.92$, $p = .03$ (η^2 =. 03). Also, differences were found with the variable 'academic level' $F(1, 147) = 8.15$, $p = .01$ ($\eta^2 = .06$) and the age covariate $F(1, 147) = 6.70$, $p = .01$ ($\eta^2 = .05$). In contrast, no significant differences were found regarding the variable 'learning environment'. It is important to point out that the highest means for the maze task resolution were found in setting 1 among 6th grade children (M = 8.40) and in setting 2 among 5th grade children (M = 8.31) and 6th grade children (M = 8.42) (see Table 5).

An ANCOVA with three fixed-effect factors (gender, learning environment and academic level) was used to compare RQ2, with age as a covariate. With respect to the gender and environment variables, no significant differences were found in student satisfaction with the execution of Blockly Games tasks. Regarding the academic level variable, differences were found in satisfaction, specifically in item 3 which referred to the student's motivation to continue performing this type of task $F(1, 147) = 8.15$, $p = .01$ ($\eta^2 = 0.06$). The highest Means were found in setting 1 among boys in 5th grade (M = 5.00) and 6th grade (M = 4.90) and among girls in 6th grade (M = 5.00) (see Table 5).

Discussion

Virtual Learning Environments (VLEs), specifically those using the serious games technique (Liu & Liu, 2020) implemented in conjunction with SRL (Taub et al., 2014) facilitate self-regulated instruction and increase student performance (Cloude, 2019; Veenman et al., 2014). The use of this methodology has been shown to be very effective for the instruction of STEM subjects (Alsawaier, 2018; Bovermann & Bastiaens, 2020; Dreimane, 2019; Queiruga et al., 2020; Sáiz-Manzanares & Marticorena-Sánchez, 2016; Zainal, 2018). In line with the conclusions of these investigations, the results of this study indicate that the performance of students in tasks of initiation to computer programming, in which both serious games (Blockly Games) and

Table 2
Example of a self-regulating dialogue in the resolution of the Blockly Games maze task on level 1

<table>
<tr><th>Teacher</th><th>Task resolution strategies</th><th>Dialogues, task resolution modeling</th></tr>
<tr><td>Defining task 1</td><td>“You have to choose the avatar of your choice (the Panda Bear, the astronaut or the person) reach the goal.”</td><td>
</td></tr>
<tr><td>Resolution-oriented strategies</td><td>“What do we have to do?”</td><td>“In this task the panda bear has to go straight on to reach the goal.”</td></tr>
<tr><td>Task resolution planning strategies</td><td>“How are we going to do it?”</td><td>“For the panda bear to go straight on, you just have to use the forward block a couple of times. As you can see, we take the block and join it twice, then press run program and the avatar reaches the goal. Bravo, we have solved the first maze.”
“Let’s start”</td></tr>
<tr><td>Resolution process evaluation strategies</td><td>“How are we solving the maze?”</td><td>
</td></tr>
<tr><td colspan="3">
</td></tr>
<tr><td></td><td>“We have to learn the code”</td><td>“Look this is the code, it just means we move forward the panda bear and we have to do it twice”
moveForward();
moveForward();</td></tr>
</table>

SRL are applied, has been very satisfactory among both boys and girls (Mean of 8 out of 10). However, performance was higher in boys. These results, on the one hand, support the theory that the combined use of SRL with serious games increases students' motivation to approach these subjects through play (Cerezo et al., 2010; Contreras et al., 2019; Díaz-Lauzurica & Moreno-Salinas, 2019; Taub et al., 2014; Zheng et al., 2020; Zimmerman & Moylan, 2009). The possible causes are that this methodology implements the use of simulation (Patti, 2019) together with the use of the figure of an avatar (Taub et al., 2014) that provides feedback on the execution, which facilitates an avoidance of the fear of error (Brooks et al., 2019; Coertjens, 2018; Conteras et al., 2019; Hattie & Clarke, 2018; Park et al., 2019) from a graduation of difficulty in the presentation of tasks (Cherry et al., 2019; Kretschmer & Terharen, 2019; Patti, 2019; Sáiz et al., 2019). However, this study found a difference in performance in Blockly Games in favor of the male gender, which supports the findings of the studies by García-Holgado et al. (2019), indicating that gender differences begin before secondary education. The degree of satisfaction of students with the performance of tasks at the beginning of the programming in Blockly Games yielded very satisfactory results among boys and girls, however more motivation to continue working on similar activities was observed among boys rather than girls.

On the other hand, it seems that the place where the activity is implemented is not significant in terms of performance or student satisfaction. Where differences have been found is with respect to the academic level variable and the age covariate. This aspect can be explained from the hypothesis that up to the age of 11, boys and girls would have a similar behaviour towards the execution of STEM tasks (Bokova, 2018). However, at the age of 11, this difference may already be starting to appear (it should be taken into account that in 6th grade of Primary Education there are students who are older than 11). This coincides with the findings of studies by Verdugo-Castro et al., (2019) that the gender gap already begins in Primary Education. It leads to a necessary reflection on the part of educationalists on what can happen for this change to take place; possible explanations could be related to the influence of variables such as: the learning history of students, the teaching style of the teacher, the upbringing style of families or other factors related to the social and/or family environment of boys and girls (Brooks et al., 2019; Gallego-Durán et al., 2014). Therefore, future research will be aimed at studying the effect of these variables and checking whether there are differences between boys and girls.

Finally, although the results found in this study are in line with those found in the research that has supported it, these results should be treated with caution, taking into account the characteristics of the sample and the time of instruction applied.

In summary, more research is needed on this topic that provides representative data when conducted in natural learning contexts (Girvan & Savage, 2019; Queiruga et al., 2020; Sáiz-Manzanares et al., 2019). The ultimate goal is for them to serve educational leaders in governments and institutions as a basis for decision-making on the design of instruction in STEM subjects at each stage of the education system (Stehle & Peters-Burton, 2019), in order to overcome the gender gap and ensure digital literacy in society. In this regard, it is important to note that STEM subjects, specifically those related to learning computer programming in the stages prior to Secondary Education, are not explicitly part of the curricula. It is a fact highlighted in many research projects (Conde-González et al., 2019; García-Holgado, Camacho-Díaz et al., 2019; García-Peñalvo, 2019), upon the premise of enhancing STEM teaching in regular teaching-learning environments, since the society of the 21st century requires digital literacy for the entire population (Azevedo & Gašević, 2019; Cloude et al., 2019; Taub et al., 2018) that must begin in the initial years of schooling (Stehle & Peters-Burton, 2019).

Table 3

Student satisfaction survey with Blockly games maze workshop, adapted from Appianing and Van Eck (2018)

NAME:	SURNAME:				
ID CODE:	AGE:				
GENDER:	COURSE:				
LEVEL:	CENTER:				
Closed responses are measured on a Likert-type scale of 1 to 5 where 1 is not at all, never or nothing and 5 is always, fully or strongly, in agreement.					
Gradation scale	**1**	**2**	**3**	**4**	**5**
1. I had already done similar activities.	1	2	3	4	5
2. I found the activity interesting.	1	2	3	4	5
3. I would like to do more of these activities.	1	2	3	4	5
4. I understood what had to be done in the activity.	1	2	3	4	5
5. I liked the activity because I worked with other colleagues.	1	2	3	4	5

Acknowledgements

This research was funded by Consejería de Educación de la Junta de Castilla y León through the Dirección General de Innovación y Equidad Educativa (Spain), grant number EDUCYL2018_04 under project «Igualdad de género en materias científico-tecnológicas: análisis comparativo estudiantes STEM». The Dirección Provincial de Educación de Burgos, and the Direction of educational establishments C.P. "Antonio Machado" and C.P. "Francisco de Vitoria" (Spain) all participated in this project. We are also thankful for the families that consented to the participation of their children in this study.

Table 4

Indicators of asymmetry and kurtosis in the pre-knowledge variable

	N	Minimum	Maximum	M	SD	Asymmetry	SEA	Kurtosis	SEK
Previous knowledge	147	1	5	2.66	1.69	0.33	0.20	-1.61	0.40

Note: *N* = number of students; *M* = Mean; *SD* = Standard Deviation; *SEA* = Standard Error Asymmetry; *SEK* = Standard Error Kurtosis

Table 5
ANCOVA with three fixed-effect variables (gender, learning environment, and academic level) and covariate age

	Environment 1 n = 43				**Environment 2 n = 104**						
	Child n = 20		**Girl n = 23**		**Child n = 56**		**Girl n = 48**		**F(1,147)**	**p**	**η^2**
	5º	**6º**	**5º**	**6º**	**5º**	**6º**	**5º**	**6º**			
	n = 10	**n = 10**	**n = 11**	**n = 12**	**n = 32**	**n = 24**	**n = 32**	**n = 16**			
	M(SD)	**M(SD)**	**M(SD)**	**M(SD)**	**M(SD)**	**M(SD)**	**M(SD)**	**M(SD)**			
Independent variable gender											
Results in Blocking Games	8.00(1.41)	8.40(1.50)	7.18(1.60)	8.25(1.29)	8.31(1.44)	8.42(1.31)	7.59(1.81)	7.75(1.61)	4.92	. 03*	. 03
2	4.10(1.66)	4.60(.70)	4.55(1.24)	4.75(.62)	4.59(.84)	4.67(.57)	4.81(.74)	4.63(.74)	1.38	. 24	. 01
3	5.00(.00)	4.90(.32)	4.64(1.21)	5.00(.00)	4.75(. 51)	4.83(.64)	4.47(1.22)	4.63(.62)	2.16	. 14	. 02
4	4.40(1.08)	4.50(.97)	4.27(1.27)	4.58(.67)	4.78(.49)	4.75(.53)	4.63(.79)	4.69(.60)	0.25	. 62	. 002
5	4.70(.68)	4.50(1.27)	4.36(1.30)	4.25(1.21)	4.56(.91)	3.92(1.25)	4.78(.55)	4.44(1.03)	0.03	. 86	. 00
Environmentally independent variable											
Results in Blocking Games									0.14	. 72	. 00
2									150	. 22	. 01
3									1.81	. 18	. 01
4									4.02	. 05	. 03
5									0.01	. 93	. 00
Independent variable level											
Results in Blocking Games									8.15	. 01*	. 06
2									3.21	. 08	. 02
3									7.35	. 01*	. 05
4									1.23	. 27	. 01
5									0.51	. 48	. 004
Covariate age											
Results in Blocking Games									6.70	.01*	.05
1									2.79	.10	.02
3									9.93	.002	.07
4									0.60	.44	.004
5									1.05	.31	.01

Note: *M* = Mean; *SD* = Standard deviation; 2 = I found the activity interesting; 3 = I would like to do more activities of this type; 4 = I understood what to do in the activity; 5 = I liked the activity, because I worked with other colleagues.
* p < .05

References

Appianing, J., & Van Eck, R. N. (2018). Development and validation of the Value-Expectancy STEM Assessment Scale for students in higher education. *International Journal of STEM Education*, *5*(24), 1-16. https://doi.org/10.1186/s40594-018-0121-8

Alsawaier, R. S. (2018). The effect of gamification on motivation and engagement. *International Journal of Information and Learning Technology*, *35*(1), 56-79. https://doi.org/10.1108/IJILT-02-2017-0009

Azevedo, R., & Gašević, D. (2019). Analyzing Multimodal Multichannel Data about Self-Regulated Learning with Advanced Learning Technologies: Issues and Challenges. *Computers in Human Behavior*, *96*, 207-210. https://doi.org/10.1016/j.chb.2019.03.025

Bandalos, D. L., & Finney, S. J. (2001). Item parceling issues in structural equation modeling. *New Developments and Techniques in Structural Equation Modeling*. Lawrence Erlbaum Associates Publishers.

Blockly Games (2020, May 23). Google's Blockly Games is a series of educational games that teach programming. It is based on the Blockly library. All code is free and open source. https://www.blockly.games/

Bokova, I. G. (2018). *Cracking the code: Girls' and women's education in science, technology, enginering and mathematics (STEM)*. United Nations Educational, Scientific and Cultural Organization

Bovermann, K., & Bastiaens, T.J. (2020). Towards a motivational design? Connecting gamification user types and online learning activities.

Research and Practice in *Technology Enhanced Learning*, *15*(1), 1-18. https://doi.org/10.1186/s41039-019-0121-4

Brooks, C., Carroll, A., Gillies, R. M., & Hattie, J. (2019). A matrix of feedback for learning. *Australian Journal of Teacher Education*, *44*(4), 14-32. https://doi.org/10.14221/ajte.2018v44n4.2

Clark, C. A. (1970). *Serious Games*. University Press of America.

Cloude, E. B., Taub, M., Lester, J., & Azevedo, R. (2019). The Role of Achievement Goal Orientation on Metacognitive Process Use in Game-Based Learning. In S. Isotani et al. (Eds.), *Artificial Intelligence in Education* (pp. 36-40). Springer International Publishing. https://doi.org/10.1007/978-3-030-23207-8_7

Coertjens, L. (2018). The relation between cognitive and metacognitive processing: Building bridges between the SRL, MDL, and SAL domains. *British Journal of Educational Psychology*, *88*(1), 138-151. https://doi.org/10.1111/bjep.12214

Contreras, M., Bonilla, D., & Peña Pérez Negrón, A. (2019). Proposal of a Model for the Development of Labor Competencies Based on Serious Games in the Context of Industry 4.0. In J. Mejia, M. Muñoz, Á. Rocha, A. Peña & M. Pérez-Cisneros (Eds.), *Trends and Applications in Software Engineering* (pp. 80-87). Springer International Publishing.

Cerezo, R., Calderón, V., & Romero, C. (2019). A holographic mobile-based application for practicing pronunciation of basic English vocabulary for Spanish speaking children. *International Journal of Human-Computer Studies*, *124*, 13-25. https://doi.org/10.1016/j.ijhcs.2018.11.009

Cerezo, R., Núñez, J.C., Rosário, P., Valle, A., Rodríguez, S., & Bernardo, A. (2010). New media for the promotion of self-regulated learning in higher education. *Psicothema*, *22*(2), 306-315. http://www.psicothema.com/pdf/3731.pdf

Conde-González, M. Á., Fernández-Llamas, C., Alves, J. F., Ramos, M. J., Celis Tena, S., Gonçalves, J., Lima, J., Reimann, D., Jormanainen, I., & García-Peñalvo, F. J. (2019). RoboSTEAM - A Challenge Based Learning Approach for integrating STEAM and develop Computational Thinking. In M. Á. Conde-González, F. J. Rodríguez-Sedano, C. Fernández-Llamas, & F. J. García-Peñalvo (Eds.), *Proceedings of TEEM'19: Technological Ecosystems for Enhancing Multiculturality 2019*, *Leon, Spain,* (pp. 24-30). Association for Computing Machinery. https://doi.org/10.1145/3362789.3362893

Díaz-Lauzurica, B., & Moreno-Salinas, D. (2019). Computational thinking and robotics: A teaching experience in compulsory secondary education with students with high degree of apathy and demotivation. *Sustainability (Switzerland)*, *11*(5109), 1-21. https://doi.org/10.3390/su11185109

Dreimane, S. (2019). Gamification for education: Review of current publications. In L. Daniela (Ed.), *Didactics of smart pedagogy* (pp. 453-464). Springer. https://doi.org/10.1007/978-3-030-01551-0_23

Gallego-Durán, F. J., Villagrá-Arnedo, C.-J., Satorre Cuerda, R., Compañia, P., Molina-Carmona, R., & Llorens Largo, F. (2014). Overview: Serious games, gamification and much more. *ReVision*, *7*(2), 13-23. http://hdl.handle.net/10045/37972

García-Holgado, A., Camacho-Díaz, A., & García-Peñalvo, F. J. (2019). Engaging women into STEM in Latin America: W-STEM project. In M. Á. Conde-González, F. J. Rodríguez-Sedano, C. Fernández-Llamas, & F. J. García-Peñalvo (Eds.), *Proceedings of TEEM'19: Technological Ecosystems for Enhancing Multiculturality 2019*, *Leon, Spain* (pp. 232-239). Association for Computing Machinery. https://doi.org/10.1145/3362789.3362902

García-Holgado, A., Verdugo-Castro, S., Sánchez-Gómez, M. C., & García-Peñalvo, F.J. (2019). Trends in studies developed in Europe focused on the gender gap in STEM. In M. Á. Conde-González, F. J. Rodríguez-Sedano, C. Fernández-Llamas, & F. J. García-Peñalvo (Eds.), *Proceedings of the XX International Conference on Human Computer Interaction Article Nº 47* (pp. 1-8). Association for Computing Machinery. https://doi.org/10.1145/3335595.3335607

García-Peñalvo, F. J., Bello, A., Domínguez, A., & Romero Chacón, R. M. (2019). Gender Balance Actions, Policies and Strategies for STEM: Results from a World Cafe Conversation. *Education in the Knowledge Society*, *20*, 1-31. https://doi.org/10.14201/eks2019_20_a31

Girvan, C., & Savage, T. (2019). Virtual worlds: A new environment for constructionist learning. *Computers in Human Behavior*, *99*, 396-414. https://doi.org/10.1016/j.chb.2019.03.017

Hattie, J., & Clarke, S. (2018). *Visible learning: Feedback. Visible Learning: Feedback*. Routledge. https://doi.org/10.4324/9780429485480

Kintsakis, D., & Rangoussi, M. (2017, April 26-30). An early introduction to STEM education: Teaching computer programming principles to 5th graders through an e-learning platform: A game-based approach. In *IEEE Global Engineering Education Conference (EDUCON) Athens, Greece* (pp. 17-23). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/EDUCON.2017.7942816.

Kretschmer, V., & Terharen, A. (2019). Serious Games in Virtual Environments: Cognitive Ergonomic Trainings for Workplaces in Intralogistics. In T. Z. Ahram (Ed.), *Advances in Human Factors in Wearable Technologies and Game Design, 795* (pp. 266-274). Springer International Publishing. https://doi.org/10.1007/978-3-319-94619-1_26

Liu, S., & Liu, M. (2020). The impact of learner metacognition and goal orientation on problem-solving in a serious game environment. *Computers in Human Behavior*, *102*, 151-165 https://doi.org/10.1016/j.chb.2019.08.021

Lodge, J. M., Kennedy, G., & Hattie, J. (2018). Understanding, assessing and enhancing student evaluative judgement in digital environments. In D. Boud, R. Ajjawi, P. Dawson & J. Tai (Eds.), *Developing Evaluative Judgement in Higher Education: Assessment for Knowing and Producing Quality Work* (pp. 70-78). https://doi.org/10.4324/9781315109251

Marji, M. (2014). *Learn to program with scratch: A visual introduction to Programming with Games, Art, Science and Math*. No starch Press.

Nogueira-Frazão, A. G., & Martínez-Solana, Y. (2019). Digital Infographics: The Key to Information Paradigm. In M. Túñez-López, V.-A. Martínez-Fernández, X. López-García, X. Rúas-Araújo, & F. Campos-Freire (Eds.), *Communication: Innovation and Quality* (pp. 293-312). Springer International Publishing. https://doi.org/10.1007/978-3-319-91860-0_17

Park, C. L., Williams, M. K., Hernández, P. R., Agocha, V. B., Carney, L. M., DePetris, A. E., & Lee, S. Y. (2019). Self-regulation and STEM persistence in minority and non-minority students across the first year of college. *Social Psychology of Education*, *22*(1), 91-112. https://doi.org/10.1007/s11218-018-9465-7

Patti, I. (2019). Serious Games as Creative Tools to Approach Design. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Eds.), *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)* (pp. 1995-2007), Florence, Italy. Springer International Publishing. https://doi.org/10.1007/978-3-319-96071-5_210

Queiruga-Dios, M.A., López-Iñesta, Díez-Ojeda, M., Sáiz-Manzanares, M.C., & Vázquez-Dorrío, B. (2020). Citizen science for scientific literacy and the attainment of SDGs in formal education. *Sustainability*, *12*(4283), 1-18. https://doi.org/10.3390/su12104283

RoboSTEM (February, 2020). Proyecto europeo de innovación y apoyo escolar [European project on innovation and school support]. http://robosteamproject.eu/.

Sáiz-Manzanares, M.C., & Marticorena-Sánchez, R. (2016). Metacognition, Self-Regulation and Feedback for Object-Oriented Programming Problem-Solving. In J. Benson. *Metacognition: Theory, performance and current research* (pp. 43-94). Nova Publishers.

Sáiz-Manzanares, M.C., Queiruga-Dios, M.A., García-Osorio, C.I., Montero-García, E., & Rodríguez-Medina, J. (2019). Observation of Metacognitive Skills in Natural Environments: A Longitudinal Study With Mixed Methods. *Frontiers in Psychology, 10*(2398), 1-13. https://doi.org/10.3389/fpsyg.2019.02398

SPSS (2016). Statistical Package for the Social Sciences (SPSS) (Version 24). IBM.

Stehle, S. M., & Peters-Burton, E. E. (2019). Developing student 21st Century skills in selected exemplary inclusive STEM high schools. *International Journal of STEM Education*, *6*(39), 1-15. https://doi.org/10.1186/s40594-019-0192-1

Taub, M., Azevedo, R., Bradbury, A. E., Millar, G. C., & Lester, J. (2018). Using sequence mining to reveal the efficiency in scientific reasoning during STEM learning with a game-based learning environment. *Learning and Instruction*, *54*, 93-103. https://doi.org/10.1016/j.learninstruc.2017.08.005

Taub, M., Azevedo, R., Mudrick, N., Clodfelter, E., & Bouchet, F. (2014, June 23-27). Can Scaffolds from Pedagogical Agents Influence Effective Completion of Sub-Goals during Learning with a Multi-Agent Hypermedia-Learning Environment? *Learning and becoming in practice: The International Conference of the Learning Sciences (ICLS) 2014, Boulder, CO, United States*.

Troiano, G. M., Snodgrass, S., Argimak, E., Robles, G., Smith, G., … Harteveld, C. (2019). Is My Game OK Dr. Scratch? Exploring Programming and Computational Thinking Development via Metrics in Student-Designed Serious Games for STEM. In J. A. Fails (Ed.), *IDC '19: Proceedings of the 18th ACM International Conference on Interaction Design and Children* (pp. 208-219). Association for Computing Machinery. https://doi.org/10.1145/3311927.3323152.

Tucker-Raymond, E., Puttick, G., Cassidy, M., Harteveld, C., & Troiano, G. M. (2019). " I Broke Your Game ! ": Critique among middle schoolers designing computer games about climate change. *International Journal of SEM Education*, *6*(41), 1-16. https://doi.org/10.1186/s40594-019-0194-z

Veenman, M. V. J., Bavelaar, L., De Wolf, L., & Van Haaren, M. G. P. (2014). The on-line assessment of metacognitive skills in a computerized learning environment. *Learning and Individual Differences*, *29*, 123-130. https://doi.org/10.1016/j.lindif.2013.01.003

W-STEM (January, 2019). *European Project of Engaging Women into STEM: Building the future of Latin America*. https://wstemproject.eu/es/inicio/.

Zainal, N. F. A., Din, R., Majid, N. A. A., Nasrudin, M. F., & Rahman, A. H. A. (2018). Primary and secondary school students perspective on Kolb-based STEM module and robotic prototype. *International Journal on Advanced Science, Engineering and Information Technology*, *8*(4-2), 1394-1401. https://doi.org/10.18517/ijaseit.8.4-2.6794

Zheng, J., Xing, W., Zhu, G., Chen, G., Zhao, H., & Xie, C. (2020). Profiling self-regulation behaviors in STEM learning of engineering design. *Computers & Education*, *143*(103669), 1-13. https://doi.org/10.1016/j.compedu.2019.103669

Zimmerman, B. J., & Moylan, A. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 299-315). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9780203876428