Investigating Video Classes Formats for Teaching Digital Game Programming in High School

Thiago Reis da Silva[1], Alan de Oliveira Santana[2], Eduardo Henrique da Silva Aranha[2]

[1]Instituto Federal de Educação, Ciência e Tecnologia do Maranhão – IFMA
Campus São João dos Patos – MA – Brasil

[2]Programa de Pós-Graduação em Sistemas Computacionais - PPgSC
Universidade Federal do Rio Grande Norte – UFRN

thiago.reis@ifma.edu.br, alandeoliveirasantana@gmail.com,
eduardoaranha@dimap.ufrn.br

Abstract. *[Context] In this study we conducted a controlled experiment comparing two formats of digital game programming teaching classes - Virtual Tutor Format (VTF) and Videotaped Record Format (FVG). [Objective] This paper compares the influence of these digital game programming teaching formats. [Method] was conducted a controlled experiment with 20 students of the first year of high school, using the experimental plan Quadrado Latino. [Results] Although both formats proved adequate in the process of learning digital game programming, VTF obtained the best results, being the preferred format for 70% of the participating students. [Conclusions] The results suggest that the use of VTF is promising for the acquisition of digital game programming skills when programmers have little or no practical programming experience.*

Resumo. *[Contexto] neste estudo realizamos um experimento controlado comparando dois formatos de aulas de ensino de programação de jogos digitais – Formato Tutor Virtual (FTV) e Formato Videoaula Gravada (FVG). [Objetivo] este trabalho compara a influência desses formatos de ensino de programação de jogos digitais. [Método] foi realizado um experimento controlado com 20 alunos do primeiro ano do ensino médio, utilizando o plano experimental Quadrado Latino. [Resultados] embora ambos os formatos mostraram-se adequados no processo de aprendizagem de programação de jogos digitais, o FTV obteve os melhores resultados, sendo o formato preferido por 70% dos alunos participantes. [Conclusões] os resultados sugerem que o uso do FTV é promissor para a aquisição de habilidades de programação de jogos digitais, quando os programadores têm pouca ou nenhuma experiência prática com programação.*

1. Introdução

The demand for qualified computer professionals is a crescent, thus increasing the need to: (i) more available vacancies at computer courses; (ii) reduce the dropout rates of these courses. In that sense, computer courses need to attract more students and prepare them better for the needs of the XXI century [ACM 2005]. This vocational intervention

may occur before the student getting into the university, still in high school. For this, it is important to present computing interestingly, motivating and challenging them.

However, although the nature of computing and student generation had changed in the last years, the teaching programming still is done most times in traditional ways that may not be adequate to keep up with evolutions [Barnes *et. al.* 2008].

One approach to engaging students in computing courses in an interesting way is the use of digital games [Jesus e Raabe 2010]. Digital games are being considered as powerful educational tools, and it is believed that their use results in a wide range of benefits. Their use increases the effectiveness of the learning process and also the interest, motivation, and persistence of the student [Silva *et. al.* 2016].

To reach high school students, another complementary method used is the video class. In Fuentes, Losada e Martin (2014), the authors report that video-based teaching method can improve students' ability to analyze and solve computational problems. Despite the important role that video classes have been playing in education, there are still few empirical studies on its use in high school, especially in the teaching of digital game programming. This article intends to act in this gap and to produce results so that the theoretical discussions and the practical application of video classes can be better grounded and directed in the teaching digital game programming context.

In this sense, this article presents the planning and results of a controlled experiment about two class formats for teaching digital game programming. Results showed that both analyzed formats presented good performances, but there are differences between them regarding efficiency and engagement.

2. Context of Study

This study was carried out with 20 students from the 1st year of a high school and integrated technical school. The selection of students was based on an invitation to the morning and afternoon shift students who were interested in participating in a study related to learning how to develop digital games. These students used two formats for the development of digital games: VTF and RVF.

The RVF follows a step by step model, which presents a game to be developed and the complete step by step for its development. Following the steps indicated by the video, in the end, the student should be able to develop the game. This format follows the style of so-called tutorials and the videos created are composed of a recording of 20 to 65 minutes, depending on the game presented in class. In this format, in previous realized studies [Silva e Aranha 2016], it was detected that the longer classes become tiresome and that they end up not being very challenging, for showing the whole step by step. Faced with this problem, the VTF was developed [Silva *et. al.* 2015].

The idea of VTF is that it be presented as a set of activities to teach game programming to elementary school students mediated by video classes and a Virtual Tutor that leads the student. The objective is to act as mediator and guide of the student during the development of a game, objecting to establish an effective and attractive learning process. This format uses the video classes in a format called Missions, following a puzzle style, which presents an opening video showing the game to be developed and challenging the student to create it.

And then, instead of following a step by step that at the end of the video has the game developed (RVF), the student will have access to a set of other generic videos. These videos are resources videos to develop the game, presenting the step by step to implement certain characteristics or functionalities of a game (ex: how to insert background images, how to give movement to characters of the game, among others). The resources videos present these characteristics in a generic, game-independent fashion, and are reusable for all games that require the features or functionality to be implemented. For example, if game A and game B have elements that shoot, both classes of these games will refer to the same Bullet behavior resource video of the Construct2 tool.

3. Planining of Experiment

For the planning and the realization of the controlled experiment presented in this work, we followed the process indicated by [Wohlin *et. al.* 2012], which includes defining the general objective, research questions, the experimental design and the analysis procedure.

3.1. Objective and Research Question

The objective of this study is delineated as from the Goal/ Question/Metric (GQM) template [Basili 1996]: to analyze the formats (Virtual Tutor Format (VTF) and Recorded Video Classes Format (RVF) in teaching digital game programming, with the purpose of evaluating and comparing the formats, with respect to the student's effort, learning and preference, from the point of view of the students and in the context of students of the first year of high school.

To achieve this goal, the following Research Questions (RQ) have been defined:

- **RQ1:** Which of the formats analyzed contribute most to student engagement?

- **RQ2:** What format promotes more motivating and stimulating learning?

- **RQ3:** In what format do students need more time to watch video classes and develop games?

To answer these RQs, a questionnaire was defined to capture information about Motivation, Satisfaction, Stimulation and Confidence of students throughout the study. This questionnaire was adapted from the study of [Savi, Wangenheim e Borgatto 2010]. The dimensions and respective items that we adapted are listed in Table 1. The study was performed in a context different from that proposed by [Savi, Wangenheim e Borgatto 2010], but with the same objective to evaluate Motivation, Satisfaction, Stimulation and Confidence. Thus, to respond to RQ1, the items Confidence and Satisfaction were analyzed; For RQ2, the items Stimulus and Motivation; And for RQ3, the time that students spent to develop the games in the analyzed formats.

The questionnaire was composed of: (i) two assertions regarding Confidence, Satisfaction, Stimulus and Motivation; (ii) five dimensions in a Likert scale for each one of the alternatives of answer: Strongly Agree (SA), Agree (A), Neutral (N), Disagree (D) and Strongly Disagree (SD).

Table 1: Aspects Assessed by the Questionnaire. Source: Adapted From [Savi, Wangenheim e Borgatto 2010].

Category	Dimension
Confidence	Using VTF/RVF made me confident that I was actually learning
Satisfaction	I was pleased to watch the video class and develop the game on VTF/RVF
Stimulus	I felt stimulated to learn from the VTF/RVF
Motivation	Experience with game development in VTF/RVF motivates me to learn game programming

3.2. Experimental Material

The proposed study aimed to apply the digital game programming concepts for first-year high school students using the Construct2 tool, in two game programming teaching Formats with using video classes (RVF and VTF). The decision to use Construct2 was the low learning curve, which enables the production of good results in a short space of time, generating executable versions accessible from freely available Web browsers. And also for being the game development tool adopted in the research laboratory of the researchers involved.

In order to execute the study, two classes were elaborated, one aiming to teach the development of the Arkanoid game and the other of a Nave game. The Arkanoid game consists of a ball, a racket and colored blocks, the goal of which is to destroy all colored blocks with the ball. And the Nave game aims to control an aircraft to shoot and destroy enemy aircraft that try to eliminate it, gaining points whenever it destroys an enemy aircraft. For both games (Arkanoid and Nave) were recorded classes in the two formats analyzed. The Arkanoid game has 14 videos, each video is in between 28 seconds and 6 minutes.

3.3. Experimental Design

Defining an appropriate design for the execution of the experiment is important because it optimizes the use of resources by considering the three basic principles of experiment design: replication, blocking, and randomization [Wohlin *et. al.* 2012]. The blocking principle indicates that we need to control factors that may influence the response variables. For the case of this study, there are two factors that need to have their effects controlled. The first factor is the individual experience and motivation of each student participating in the experiment, which can contribute both positively and negatively to the study results. The second factor are the classes (games), whose complexity and attractiveness can also have a significant effect on the results.

In this context, we chose a design in Latin Square [Box, Hunter e Hunter 2005]. This was choice because this design to allow the control of two factors involved in the experiment [Leroy 2011]. In an experiment of Latin Square type, the controlled factors and the treatments are organized into quadratic matrices two by two (two rows and two columns), as shown in Table 2.

To apply this project to our experiment, we position the students as indexes of the lines, and the games as indexes of the columns of each Latin Square (Table 2). Applying the design of the Latin Square, each student experiences the two treatments

(VTF and RVF) using the two distinct classes (Arkanoid and Nave game). A random select indicates which class will be used for each treatment to be tested. To obtain replicas in the experiment, we used several replicas of squares changing only the indexes of the lines (students) and maintaining the columns (games), besides performing a new procedure of randomization about the treatment (class format) to be used by each student in each class.

Table 2: Experimental Design.

	Arkanoid Game (J1)	Nave Game (J2)		Arkanoid Game (J1)	Nave Game (J2)	…
Student1	VTF	RVF	Student 3	RVF	VTF	
Student2	RVF	VTF	Student 4	VTF	RVF	
Square 1			Square 2			

The experiment was performed in three days, each day with a four-hour session. The first day was aimed at explaining how the experiment would be and presents an overview of the Construct2 tool and how students should proceed with the development of the experiment. The main concern was to instruct students to follow development correctly. Otherwise, they could try to explore the tool for a purpose other than the focus of the study. In addition, the experiment manager recommended that students asked questions whenever they want and that they should direct their questions exclusively to the manager of the experiment, without asking questions of the other students. The second and third day were dedicated to the development of the games, each day with one of the formats, according as determined by the randomization of the experiment and monitored by the manager of the experiment. At the end of each day, the students answered a questionnaire.

3.4. Data Analysis Procedures

To respond to RQ1 and RQ2, we analyzed the responses (Likert scale) with respect to the items identified in Table 1. The analysis begins with the visualization of a bar graph to identify the percentages of response frequency for each item in the scale Likert. After that, the responses were converted to a scale of 1 to 5 and the sample mean was calculated. Statistical tests were used to verify the existence of statistically significant differences with a level of significance of 0.05. The RQ3 is answered based on the analysis of the time students spent to develop the games in each lesson format investigated. For reasons of space, graphs, figures and tables about the data collected could not be presented in this paper.

For data analysis, were used the R programming language and the RStudio and Minitab software. It was verified the supposition of normality for the sample data (time) with the Anderson-Darling test. As the non-normality of the data was observed, we used the non-parametric Wilcox test.

4. Results

Following, we describe the results achieved in each RQ.

RQ1: Which of the analyzed formats contributes most to student engagement?

To answer this RQ, we analyzed the items Confidence (confidence in learning) and Satisfaction (satisfaction with the lesson) of Table 1. In the affirmative regarding

Confidence, we can observe a slight increase in the level of agreement for the VTF, obtaining 50% for Strongly Agree (SA) and 50% for Agree (A), while RVF obtained 30% for Strongly Agree (SA), 65% for Agree (A) and 5% for Neutral (N), as shown in Figure 1 (a).

(a) **(b)**
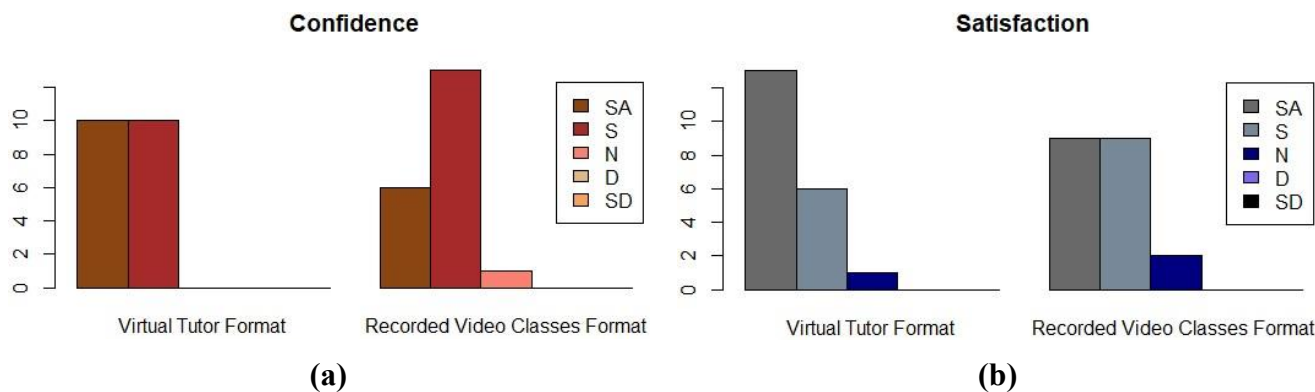
Figure 1: Bar graph of Confidence (a) and Satisfaction items (b).

In Figure 1(b), we can observe that the VTF also had a slight majority for the Satisfaction question, obtaining 65% for Strongly Agree (SA), 30% for Agree (A) and 5% for Neutral (N) while the RVF had, respectively, for Strongly Agree (SA) and Agree (A) 45% and, 10% for Neutral (N).

When converting the answers on a scale of 1 to 5, Table 3 presents the averages obtained by the participants of this study in their evaluation of Confidence and Satisfaction. As can be observed, the VTF presents an average higher than the RVF in the analysis of the two Formats.

Table 3: Average of the confidence and satisfaction items.

	VTF	RVF
Confidence	4.5	4.25
Satisfaction	4.6	4.35

It purposing to find out the statistical significance of this difference of results presented for these two items, resorted itself the Wilcox test. The resulting P-Value for the items Trust and Satisfaction are presented in Table 4. According to revealed, we cannot affirm that there is a statistically significant difference between the levels of Confidence and Satisfaction of the VTF and the RVF.

Table 4: P-Value of items.

	Confidence	Satisfaction
P-Value	0.9999	0.2355

In response to RQ1, it can be concluded that the VTF presented a slight superiority to the RVF. However, this result cannot be generalized because there was no statistically significant difference. Thus, we can assume that for the item Trust we accept the Null Hypothesis (H1.0_1): "*The average confidence level of the students who developed the game with the VTF is equal to that obtained by those who developed with the RVF*" and we reject the Alternative Hypothesis (H1.1_1). And for the item Satisfaction we also accept the Null Hypothesis (H1.0_2): "*The average level of satisfaction of the students who developed the game with the VTF is equal to that obtained by those who developed with the RVF*" and we reject the Alternative Hypothesis (H1.1_2).

RQ2: What format promotes more motivating and stimulating learning?

To support the analysis of this RQ, the Stimulus and Motivation items in Table 1 were analyzed. In the assertions regarding Stimulus, we can observe an increase in the stimulus level for VTF, where it obtained 40% for Strongly Agree (CF), 50% for Agree (A) and 10% for Neutral (N), while RVF obtained 15 % for Strongly Agree (CF), 50% for Agree (A), 15% for Neutral (N), 10% for Disagree (D) and 10% for Strongly Disagree (DF), as illustrated in Figure 2a.

However, for the Motivation item we can observe a slight superiority for the VTF, which obtained 55% for Strongly Agree (CF), 40% for Agree (A) and 5% for Neutral (N). The RVF obtained 20% for Strongly Agree (CF), 40% for Agree (A), 25% for Neutral (N), 10% for Disagreement (D) and 5% for Strongly Disagree (DF) as shown in Figure 2b.
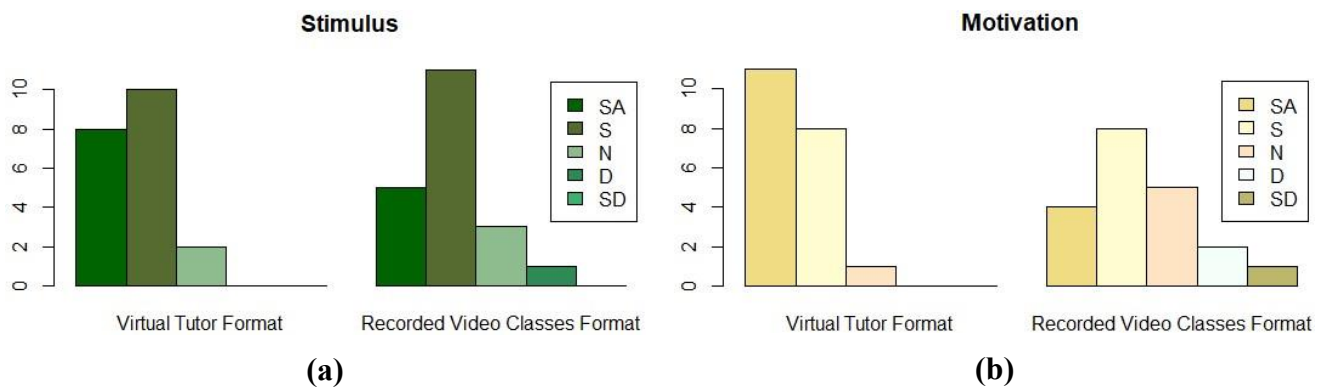
Figure 2: Bar graph of the Motivation (a) and Stimulus items (b).

The average between the formats can be observed in Table 5, presenting that the VTF had an average higher than the RVF for the items Stimulus (22.8% higher) and Motivation (25% higher).

Table 5: Average Score of Stimulus and Motivation.

	VTF	RVF
Stimulus	4.3	3.5
Motivation	4.5	3.6

With objective to verify the statistical significance of this difference of results, we used the Wilcox test again. The resulting P-value for the Motivation and Stimulus items are presented in Table 6. Thus, we have sufficient statistical evidence to state that these differences are not due to chance and that they can be generalized to similar situations.

Table 6: P-Value of the statistical tests.

	Stimulus	Motivation
P-Value	0.02146	0.04314

Therefore, we conclude that for the Stimulus item we accept the Alternative Hypothesis (H1.11): "*The average result of stimulus in learning using the VTF is different from that obtained by those who developed with the RVF*". And, for the Motivation item, we reject Null Hypothesis (H1.02) and we accept the Alternative Hypothesis (H1.12): "*The average result of stimulus in learning using the VTF is different from that obtained by those who developed with the RVF*".

RQ3: In what format do students need more time to watch video classes and develop games?

To answer this RQ we analyze the time of each student during the development of the game in the two formats of classes applied. Table 7 presents the data with the average time of the students in the formats analyzed, also the average, and general standard deviation of the formats. As a result, the students who developed the game using FVT spent an average time greater (23.6% higher) to develop the game compared to the students who developed using the RVF. In addition, the standard deviation of the students who used the VTF is relatively lower in relation to the RVF, that is, there is less variation in the students' averages in the use of VTF in relation to the RVF.

Table 7: Average of items.

Format	Average	Standard Deviation
VTF	131	34,19
RVF	106	47,71

From the collected data, a statistical test was applied in order to evaluate if there was significant difference between the results obtained by both formats in relation to the time. Thus, the Anderson-Darling normality test was initially applied to evaluate the normality of the sample. The result indicated that the data do not keep on a normal distribution for a confidence level of 95%.

Even so, we performed the Analysis of Variance (ANOVA) to perform the residue analysis. As expected the result showed a heterogeneity of the residues in relation to the adjusted values, violating the main ANOVA supposition.

In a second attempt, the data were transformed using the Box-cox method, but even then, the residue analysis showed the inapplicability of the ANOVA. Thus, it was necessary to apply a non-parametric test to evaluate the possibility of a statistically significant difference between the two formats. Then, the Wilcox test was applied with a confidence level of 95% ($\alpha = 0.05$). The test result indicated a P-value of 0.02146 lower than the confidence level adopted, confirming that there are statistically significant differences. As the average time using the VTF was higher, we accepted the Alternative Hypothesis defined.

Therefore, based on the result obtained, we reject the Null Hypothesis (H3.0$_1$): "*There is no mean time difference between the results obtained by the students who used the Virtual Tutor Format and the students who used the Video classes Recorded Format*" and we assumed the Hypothesis Alternative (H3.0$_2$): "*Students who used the RVF present results in less time than the students who used the VTF in relation to the time to develop the game*". This result shows that on average developing the game with the VTF takes about 23.6% more time

5.1. Threats To The Validity Of The Study

This section presents and discusses the threats to the validity of the study that have been identified, according to the nomenclatures defined in [Wohlin *et. al.* 2012].

- **Construction Validity**: the questionnaire used for data collection was validated by a specialist in the area. Regarding the threats, the students did not know what the research questions and hypotheses of the study, the profile data were provided voluntarily by the students themselves;

- **Conclusion Validity**: the assumptions of the statistical tests were accomplished. However, it was not possible to treat the low statistical power since the sample is considered small;

- **Internal validity**: in relation to instrumentation, the questionnaire was previously evaluated by a specialist before the students answered and the selected students did not undergo any intervention before the study;

- **External validity**: it is not possible to generalize the results since the sample is not representative of the population of basic education students of the country.

6. Conclusion

In general, it could be observed that the results report the applicability of the two Formats for student use. Nevertheless, from the results obtained it is possible to perceive that the VTF had better result. This may have occurred due to this format working more student reasoning. This result corroborates the students' preference for the format to be used, most of the students (75%) prefer to use the VTF.

While the VTF, statistically, obtained superior results for the items Stimulus and Motivation, the RVF obtained superior results for the Time item.

Other easiness identified as relevant to students, especially in relation to the VTF, is this format is developed by specific functions, so if they have an understanding of a certain content, do not need to attend such a class, already jumping to the next. It was possible to observe that the majority of the students were able to develop the games in both formats, having a preference to use the VTF. We also observed that the ease with the students could test and execute, immediately, the developed games, it helped in the learning and understanding of the worked concepts. This strategy allows students to immediately obtain feedback from the developed game.

In the study, the active participation of all the students during the execution was perceptible. Students demonstrated motivation and willingness to continue learning about game development. During class the students began to insert new elements in the game, including, new characters, fund plan, inserting new behaviors, among others. The diversity of options and possibilities of what can be created using the Construct2 Tool and in line with the Formats encourages the creation of more games and thus more and more programming learning.

The Construct2 tool has demonstrated to be suitable for teaching game programming to this audience. In terms of programming, all students were able to use commands/resources (event, behavior, and variables). Construct2 stood out for its intuitiveness, diversity of resources (scenarios and characters) available, stimulating the development of games for its ease of use.

Considering the initial results already achieved, we have the perspective of future works that may include a greater number of participants, as well as the diversification of content and game development tools for this audience.

References

ACM/IEEE-CS. (2005) "The Joint Task Force for Computing Curricula". Computing Curricula 2005. ACM/IEEE.

Barnes, T.; Powell, E.; Chaffin, A.; Lipford, H. (2008) "Game2Learn: improving the motivation of CS1 students". In: International Conference on Game Development in Computer Science Education, Miami, FL.

Basili, V. R. (1996) "Applying the goal/question/metric paradigm in the experience factory". Chapter 2 in Software Quality Assurance and Measurement: A Worldwide Perspective, Norman Fenton, International Thomson Publishing, London, UK

Box, G. E. P.; Hunter, J. S.; Hunter, W. G. (2005) "Statistics for Experimenters: Design, Innovation and Discovery". Wiley Interscience.

Jesus, E. A.; Raabe, A. L A. (2010) "Avaliação Empírica da Utilização de um Jogo para Auxiliar a Aprendizagem de Programação". In: Simpósio Brasileiro de Informática na Educação – SBIE.

Leroy, G. (2011) "Designing User Studies in Informatics". Springer.

Savi, R.; Wangenheim, C. G. V.; Borgatto, A. (2011) "Um Modelo de Avaliação de Jogos Educacionais na Engenharia de Software". In: XXV Simpósio Brasileiro de Engenharia de Software, pp.194-203, São Paulo, SP.

Silva, T. R.; Aranha, E. H. S. (2016) "An Empirical Study of Online K-12 Education for Programming Games". In: IEEE 16th International Conference on Advanced Learning Technologies (ICALT), Austin, TX, 2016, pp. 255-259.

Silva, T. R.; Aranha, E.; Fernandes, K.; Santos, F. (2016) "Um relato de experiência da aplicação de videoaulas de programação de jogos digitais para alunos da educação básica". In: XXII Workshop de Informática na Escola – WIE, Uberlândia – MG, p.141-150.

Silva, T. R.; Aranha, E. H. S.; Olivera, W.; Fernandes, K. T.; Lucena, M. J. N. R. (2015) "Investigando dois formatos de videoaulas de programação de jogos digitais para alunos do ensino médio". In: XXI Workshop de Informática na Escola – WIE, p.187-196.

Urquiza-Fuentes, J.; Hernan-Losada, I.; Martin, E. (2014) "Engaging students in creative learning tasks with social networks and videobased learning". In: IEEE Frontiers in Education Conference – FIE, pp.1-8.

Wohlin. C.; Runeson, P.; Host, M.; Ohlsson, M.; Regnell, B.; Wessl, A. (2012) "Experimentation in Software Engineering: an Introduction". Kluwer Academic Publishers.