Journal of Universal Computer Science, vol. 25, no. 7 (2019), 840-864
submitted: 16/1/19, accepted: 29/4/19, appeared: 28/7/19 © J.UCS

Effectiveness of Games in Software Project Management Education: An Experimental Study

Giani Petri
(Federal University of Santa Maria, Santa Maria/RS, Brazil
gpetri@inf.ufsm.br)

Christiane Gresse von Wangenheim
(Federal University of Santa Catarina, Florianópolis/SC, Brazil
c.wangenheim@ufsc.br)

Jean Carlo Rossa Hauck
(Federal University of Santa Catarina, Florianópolis/SC, Brazil
jean.hauck@ufsc.br)

Adriano Ferreti Borgatto
(Federal University of Santa Catarina, Florianópolis/SC, Brazil
adriano.borgatto@ufsc.br)

Abstract: Software Project Management (SPM) is considered important to ensure that software projects are delivered with success, with respect to project scope, time, cost and quality requirements. However, teaching SPM remains a challenging issue. In this context, educational games have been used in order to provide more practical opportunities in SPM education. Yet, there is a lack of studies with more rigorous research design in order to analyse the real effectiveness of such games. The objective of this study is to analyse the effectiveness of educational games with respect to the students' experience and learning in SPM education. An experimental study is conducted in two SPM courses, involving an experimental group, adopting games; and a control group, using traditional exercises. The results indicate that games provide a positive contribution to the students' perceived learning, as well as provide an engaging experience, mainly in terms of challenge, social interaction, fun and focused attention. Yet, no indication that the games provide a better learning effect than exercises could be identified, although the results show an increase in the students' performance over time playing the games. Results of this study may assist instructors to select such games and contribute to their adoption for SPM education.

Keywords: software project management, educational game, experiment, effectiveness
Categories: D.2.9, L.5.1

1 Introduction

Software Project Management (SPM) is an important knowledge area and essential to the success of software projects [PMI, 13]. It is defined as the application of project management activities to ensure that software products and services are delivered efficiently, effectively, and to the benefit of stakeholders [Bourque, 14].

SPM is taught as a sub-topic of the Software Engineering discipline following the curriculum guidelines for undergraduate computing programs provided by the Association for Computing Machinery and IEEE Computer Society [ACM, 13]. SPM concepts, methods, and principles should be approached in a rather practical than theoretical way. However, typically, SPM is taught through traditional lectures, in which students are exposed to theoretical concepts, lacking practical training in real-life scenarios [Geist, 07; Ojiako, 11; Hussein, 15].

Therefore, other instructional strategies, such as educational games, have been proposed in order to provide more practical opportunities in SPM education. Games may assist in demonstrating the application of theory and can help students to practice SPM contents in a risk-free and controlled environment [Geist, 07; Ojiako, 11; Connolly, 12; Backlund, 13]. Besides promoting entertainment, games are designed to teach certain knowledge, and/or to develop skills or attitudes [Prensky, 07; Ritterfeld, 10; Djaouti, 11]. Games used for SPM education are expected to provide several benefits, increasing learning effectiveness through a fun and engaged environment, in which students can make decisions in a project and observe the consequences, learning from these interactions [Pfahl, 01; Prensky, 07]. In this context, different educational games have been developed for SPM education, including digital games such as X-MED [Gresse von Wangenheim, 09a], SimSE [Navarro, 07], Simsoft [Bavota, 12], ProDec [Calderón, 13], as well as non-digital games, e.g., Detective Game [Gresse von Wangenheim, 14], SCRUMIA [Gresse von Wangenheim, 13b], among others. Most SPM games are simulation games, which allow to practice skills in a realistic environment while keeping the students involved. In addition, also games targeting learning objectives at lower cognitive levels are often being used as a complementary instructional strategy to revise and reinforce previously taught knowledge, e.g. through quiz games [Battistella, 16].

Although games are considered a beneficial instructional strategy for SPM education [Petri, 18b], there is a lack of well-designed studies analysing their effectiveness in teaching and learning SPM, thus, leaving their effectiveness questionable or at least not rigorously established [Calderón, 15]. Typically, existing studies analyse the contributions of one game used to teach a specific SPM content, not analysing the games' contribution over a SPM course. In addition, such evaluations of games are often carried out without a scientific rigor, lacking a clear definition of research design, data collection instruments and data analysis [Calderón, 15; Kosa, 16; Petri, 17a]. Thus, a question that arises is to which regard the expected effectiveness of games in a SPM course is real.

In order to analyse the effectiveness of educational games with respect to students' experience and learning in higher computing education, we conduct a quasi-experimental study with two groups of students: an experimental group, who adopts educational games as an active instructional strategy for SPM education; and a control group, who adopts a traditional instructional strategy using exercises. Data have been collected from both groups through pre/post-tests and evaluation questionnaires, and then analysed and discussed in order to achieve our research objective.

2 Related Work

Different studies have been conducted in order to analyse the contribution and effectiveness of games with educational purposes in different knowledge areas [Randel, 92; Virvou, 05; Gresse von Wangenheim, 09b; Backlund, 13; Capuano, 15]. In general, the results of these studies indicate that games may contribute to the improvement of the students' knowledge and also to other factors such as attention, motivation, etc. [Virvou, 05; Gresse von Wangenheim, 09b; Connolly, 12; Backlund, 13]. However, for the adoption of games in specific educational areas, more empirical studies are required in order to investigate and confirm the expected benefits and effectiveness of such games [Connolly, 12; Backlund, 13; All, 16].

In this context, the adoption of games as an active instructional strategy for SPM education has been growing and different studies have also been conducted in order to analyse their contributions and/or effectiveness [Black, 09; Petri, 17b; Rumeser, 18; [Calderón, 18a; Calderón, 19]. In this context, SPM games and its evaluations have also been analysed through systematic literature reviews [Calderón, 15; Petri, 17a; Calderón, 18b]. These systematic literature reviews report games used for SPM/computing education and analyse how the evaluations of games have been defined, in terms of goals, evaluation factors, research design, data collection instruments; executed, in terms of sample size; and analysed, in terms of data analysis methods, etc., reporting a total of eleven games for SPM education. Table 1 presents a summary of the evaluations of the games reported by the systematic literature reviews [Calderón, 15; Petri, 17a; Calderón, 18b].

The SPM games evaluated in the related studies (Table 1) are typically used in order to review and reinforce basic SPM concepts through quiz games (e.g., PM Master [Gresse von Wangenheim, 12b]) or to simulate the planning, execution, and control of a software project (e.g., SimSE [Navarro, 07], SCRUMIA [Gresse von Wangenheim, 13b], SimSoft [Bavota, 12], ProDec [Calderón, 13; Calderón, 17]). Most of them focus on a specific SPM content such as earned value management [Gresse von Wangenheim, 12a; Gresse von Wangenheim, 14], software measurement [Gresse von Wangenheim, 09a], team management [Gresse von Wangenheim, 13a], Scrum [Fernandes, 10; Gresse von Wangenheim, 13b], among others. In terms of game genres, simulation games are predominant, typically, placing the player in the control of a certain environment or activity, seeking to be as realistic as possible.

In general, the related studies provide evidence of the contributions of SPM games mainly in terms of motivation [Gresse von Wangenheim, 12a; Gresse von Wangenheim, 13a; Calderón, 17], user experience [Gresse von Wangenheim, 13b; Calderón 17], enjoyability [Fernandes, 10] and knowledge improvement/learning acquisition [Bavota, 12; Gresse von Wangenheim, 13b; Calderón, 17], based on the students' perception, in the context of a specific SPM content. However, as presented in Table 1, the evaluations of SPM games are typically conducted adopting a simple procedure (non-experimental or ad-hoc studies) with a small sample size using questionnaires for data collection, without pre- and post-tests [Calderón, 15; Petri, 17a]. Thus, these evaluation results may be limiting the identification of empirical evidence on the effectiveness of educational games in SPM education.

Game	Game genre	Research design	Evaluated factors	Instruments	Sample size	Analysis methods
Dealing with difficult people [Gresse von Wangenheim, 13a]	Simulation	Non-experimental	Learning Motivation User experience	Questionnaire	41-60	Median Frequency diagrams
DELIVER! [Gresse von Wangenheim, 12a]	Simulation	Non-experimental	Learning Motivation User experience	Questionnaire	21-40	Median Frequency diagrams
Detective Game [Gresse von Wangenheim, 14]	Simulation	Non-experimental	Learning Motivation User experience	Questionnaire	41-60	Median Frequency diagrams
Incredible Manager [Dantas, 04]	Simulation	Ad-hoc	Learning, Fun	Questionnaire	1-20	Qualitative analysis
PlayScrum [Fernandes, 10]	Simulation	Ad-hoc	Enjoyability Confidence Learning Fun	Questionnaire	1-20	Mean
PM Master [Gresse von Wangenheim, 12b]	Quiz	Non-experimental	Learning Motivation User experience	Questionnaire	21-40	Median Frequency diagrams
SCRUMIA [Gresse von Wangenheim, 13b]	Simulation	Non-experimental	Learning Motivation User experience	Questionnaire	61-80	Median Frequency diagrams
SimSE [Navarro, 07]	Simulation	Experimental	Learning	Test Interview	1-20	Mean Histogram ANOVA
Simsoft [Bavota, 12]	Simulation	Ad-hoc	Learning Satisfaction Recommendation	Questionnaire	41-60	Mean Histogram
X-MED [Gresse von Wangenheim, 09a]	Simulation	Experimental	Learning Relevance Challenge Satisfaction Fun Interest, etc.	Questionnaire Test	1-20	Mean Median SD Mann-Whitney U
ProDec [Calderón, 13; Calderón, 17]	Simulation	Non-experimental	Motivation Experience Learning acquisition	Questionnaire	1-20	Median Frequency diagrams

Table 1: Summary of the related work

Furthermore, there basically do not exist studies analysing the contributions and/or effectiveness of different games or different game genres over a SPM course. Therefore, studies with a more rigorous research design, such as experiments are required in order to analyse the real effectiveness of such games with respect to students' experience and learning over a SPM course.

3 Definition and Execution of the Study

In order to analyse the effectiveness of educational games for SPM education, an experimental study was conducted involving students from two SPM courses.

3.1 Definition

Following the Goal/Question/Metric (GQM) approach [Basili, 94], the evaluation goal is to analyse the effectiveness of educational games in students' experience and learning from the students' point of view in SPM courses in higher computing education.

In order to achieve this goal and taking in consideration the characteristics of the research context (SPM courses with a small number of students enrolled), a quasi-experimental study was chosen as research design, being the most suitable research design in our context, as randomization was impossible due to practical restrictions such as classes at different times (daytime/evening classes) [Wohlin, 12]. The study involves two groups: an experimental group, who adopts as treatment, educational games for SPM education; and a control group, who adopts as alternative treatment, online exercises in the learning management system Moodle[1]. We use a non-probability sampling technique in each group applying the convenience sampling method [Trochim, 08], in which the sample of each group is composed by all the students enrolled in each of the SPM courses in undergraduate computing programs.

Based on the evaluation goal and following the GQM approach [Basili, 94], the analysis questions (AQ) are:

AQ1: Is there a difference in students' perceived learning between the experimental and the control group?

AQ2: Is there a difference in students' experience between the experimental and the control group?

AQ3: Is there a difference in students' learning between the experimental and the control group?

AQ4: Is there a difference in students' experience between different game genres?

AQ5: Is there a difference in students' learning between different game genres?

Measurement. In order to answer these analysis questions, metrics for each analysis question are defined (Table 2 and 3). The metrics with respect to analysis question AQ1, AQ2, and AQ4 are defined based on the MEEGA+ model [Petri, 18a], a systematic, reliable and valid model for the evaluation of educational games. The model evaluates the students' perception after they played a game through a standardized questionnaire. The MEEGA+ model evaluates games with respect to player experience and perceived learning. The quality factor player experience is decomposed into dimensions of focused attention, challenge, confidence, social

[1] https://moodle.org

interaction, satisfaction, fun, and usability. Thus, adopting the MEEGA+ model, there are 17 metrics (M_1 to M_9 (Table 2) and M_{11} to M_{18} (Table 3) defined with respect to AQ1, AQ2, and AQ4 (Table 2 and 3). In order to answer the analysis questions AQ3 and AQ5, the metrics (M_{10} and M_{19}) comprise the means of the students' test scores.

Metrics		Hypothesis
Perceived Learning		
M_1	Median of students' **perceived learning** promoted by the games/exercises.	H_{0A}: There is no difference in **perceived learning** between the experimental and control group.
Students' Experience		
M_2	Median of students' perception of **usability** when using games/exercises.	H_{0B}: There is no difference in **usability** between the experimental and control group.
M_3	Median of students' perception of **confidence** when using games/exercises.	H_{0C}: There is no difference in **confidence** between the experimental and control group.
M_4	Median of students' perception of **challenge** when using games/exercises.	H_{0D}: There is no difference in **challenge** between the experimental and control group.
M_5	Median of students' perception of **satisfaction** when using games/exercises.	H_{0E}: There is no difference in **satisfaction** between the experimental and control group.
M_6	Median of students' perception of **social interaction** when using games/exercises.	H_{0F}: There is no difference in **social interaction** between the experimental and control group.
M_7	Median of students' perception of **fun** when using games/exercises.	H_{0G}: There is no difference in **fun** between the experimental and control group.
M_8	Median of students' perception of **focused attention** when using games/exercises.	H_{0H}: There is no difference in **focused attention** between the experimental and control group.
M_9	Median of students' perception of **relevance** when using games/exercises.	H_{0I}: There is no difference in **relevance** between the experimental and control group.
Learning effectiveness		
M_{10}	Mean of **students' test scores** when using games/exercises.	H_{0J}: There is no difference in **students' test scores** between the experimental and control group.

Table 2: Metrics and hypothesis to answer AQ1, AQ2, and AQ3

	Metrics	Hypothesis
Difference in students' experience per game genre		
M_{11}	Median of students' perception of **usability** when using different game genres.	**H_{0K}**: There is no difference in **usability** between different game genres.
M_{12}	Median of students' perception of **confidence** when using different game genres.	**H_{0L}**: There is no difference in **confidence** between different game genres.
M_{13}	Median of students' perception of **challenge** when using different game genres.	**H_{0M}**: There is no difference in **challenge** between different game genres.
M_{14}	Median of students' perception of **satisfaction** when using different game genres.	**H_{0N}**: There is no difference in **satisfaction** between different game genres.
M_{15}	Median of students' perception of **social interaction** when using different game genres.	**H_{0O}**: There is no difference in **social interaction** between different game genres.
M_{16}	Median of students' perception of **fun** when using different game genres.	**H_{0P}**: There is no difference in **fun** between different game genres.
M_{17}	Median of students' perception of **focused attention** when using different game genres	**H_{0Q}**: There is no difference in **focused attention** between different game genres.
M_{18}	Median of students' perception of **relevance** when using different game genres.	**H_{0R}**: There is no difference in **relevance** between different game genres.
Difference on learning effectiveness per game genre		
M_{19}	Mean of **students' test scores** when using different game genres.	**H_{0S}**: There is no difference in **students' test scores** between different game genres.

Table 3: Metrics and hypothesis to answer AQ4 and AQ5

Instrumentation. In order to collect the data, different data collection instruments are used. Metrics M_1 to M_9 and M_{11} to M_{18} are collected by adopting the standardized self-evaluation questionnaire of the MEEGA+ game evaluation model [Petri, 18a]. The MEEGA+ questionnaire has been evaluated on a large-scale in terms of reliability and validity, presenting an excellent internal consistency (Cronbach's alpha α=.927) [Petri, 18a]. It is composed of 32 items using a 5-point Likert scale ranging from -2 (strongly disagree) to 2 (strongly agree). In order to collect data with respect to M_{10} and M_{19}, a pre-test was applied in the experimental and the control group at the beginning of the SPM courses, and a post-test was conducted after the application of each game in the experimental group and after each exercise in the control group. Both pre/post-tests were designed and prepared by all authors together, with comparable content and difficulty. The tests are composed of 10-15 multiple-choice questions with respect to the knowledge approached by the respective game/exercise.

3.2 Execution

The experiment was performed in March - July 2018 in two SPM courses (INE5427 and INE5617) in undergraduate computing programs (Computer Science and Information Systems) at the Department of Informatics and Statistics of the Federal

University of Santa Catarina, Brazil. Both courses aim at teaching an understanding of key concepts and processes in software project management. Both courses have the same duration (72 hours), are similar in terms of contents (Figure 1) and occur in the same period (3rd year) of both computing programs. Demographic characteristics of the students of both SPM courses are also similar in terms of gender and age group: in the experimental group (n=52), 90% of the students are men and 92% are between 18-28 years old; in the control group (n=91), 85% are men and 91% of the students are between 18-28 years old.

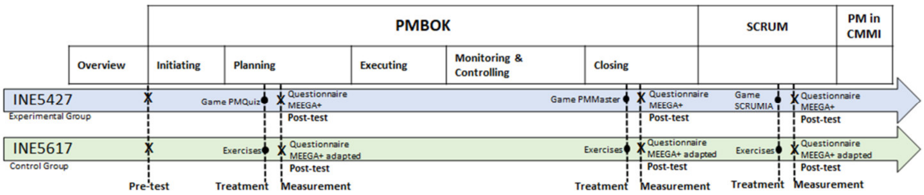


Figure 1: Measurements and sequence of contents of the SPM courses

Measurements were performed at 5 different moments during the semester. In the beginning of the SPM courses, both groups answered the same pre-test. During the semester, the course content was taught through lectures, reviewing and reinforcing the competencies through the treatments. The experimental group played three educational games (Table 4) and the control group performed three sets of online exercises (quizzes) available via Moodle, covering the same SPM contents as the games (Table 5). After each treatment (application of an educational game/exercises) the students answered the MEEGA+ questionnaire in order to collect data on metrics M_1 to M_9 and, M_{11} to M_{18} and a post-test was applied to collect data with respect to metrics M_{10} and M_{19}.

SPM content	Treatments	Session time	Participants (*n*)
Scope management	**Game 1**: **PMQuiz** [Petri, 16] is an answer-question (quiz) game to review SPM concepts, focusing on to review scope management knowledge. Players have to answer each question within a time limit using their smartphone. The player, who answer more questions correctly in the shortest time is the winner.  *Figure 2: PMQuiz game session*	30 minutes	19
PMBOK	**Game 2: PM Master** [Gresse von Wangenheim, 12b] is a (quiz) board game with questions about different PM knowledge areas, similar to Trivial Pursuit. It aims to review and reinforce basic PM concepts in accordance with the PMBOK.  *Figure 3: PM Master game session*	90 minutes	17
Scrum	**Game 3: SCRUMIA** [Gresse von Wangenheim, 13b] is a group simulation game with the purpose of planning and executing sprints of a hypothetical project by applying SCRUM. Its learning objective is to reinforce the concepts and to teach the competency to apply agile project management using SCRUM.  *Figure 4: SCRUMIA game session*	90 minutes	16

Table 4: Summary of the treatments conducted with the experimental group

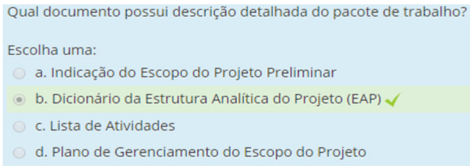
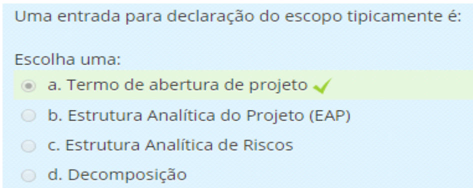
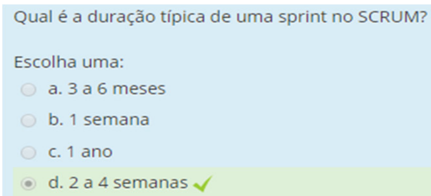
SPM content	Treatments	Session time	Participants (*n*)
Scope management	**Exercises 1:** A quiz with ten multiple-choice questions covering SPM concepts on scope management. The questions were available for the students via the learning management system (Moodle) adopted in the course. The learning objective is to review knowledge on scope management.  *Figure 5: Example of a question of the Scope management quiz*	30 minutes	29
PMBOK	**Exercises 2:** A quiz with fifteen multiple-choice questions covering PMBOK concepts from different PM knowledge areas such as scope, time, risks, and quality management. The questions were available via Moodle. The learning objective is to review knowledge on PMBOK concepts.  *Figure 6: Example of a question of the PMBOK quiz*	30 minutes	34
Scrum	**Exercises 3:** A quiz with ten multiple-choice questions on Scrum concepts. The questions were available via Moodle. The learning objective is to review Scrum concepts.  *Figure 7: Example of a question of the Scrum quiz*	30 minutes	28

Table 5: Summary of the treatments conducted with the control group

The sample size of the control and experimental groups varied during the treatments. This is due to the variation of the presence of students in the classes of the courses throughout the semester.

4 Analysis

In order to analyse the data, we used the Mann-Whitney test, an adequate test to analyse data from independent samples and with a sample size variation [Trochim, 08]. It allows a comparison of data of each group (control/experimental) and each game genre (quiz/simulation), and consequently the analysis of the questions and hypotheses [Trochim, 08]. In order to answer the questions, the test results for each metric are presented in Tables 6, 7 and 10, as well as the median and the interquartile range of the responses provided by each group. In addition, in order to analyse the learning effectiveness of the games compared to exercises (AQ3) and comparing different game genres (AQ5), we adopted the two-way analysis of variance (ANOVA), which analyses the influence of two independent variables (pre/post-test scores) on one continuous dependent variable (learning) [Trochim, 08].

AQ1: Is there a difference in students' perceived learning between the experimental and the control group?

In terms of perceived learning (M_1), students of both groups (experimental and control) identified that the activities (games and exercises) contributed in a positive way to their learning of SPM. However, the results of testing the hypothesis H_{0A}, indicate that the students of the experimental group, who played the games, had a better perception of learning than the students of the control group doing exercises (Table 6). Therefore, we can reject the null hypothesis H_{0A}.

Questionnaire item	Control Group		Experimental Group		Mann-Whitney test	
	Median	Interquartile Range	Median	Interquartile Range	p-value (2-tailed)	Test result
Perceived Learning (M1)						
The games/exercises contributed to my learning in this course.	1	0	1	1	0.022	Reject H_{0A}
The games/exercises allowed for efficient learning compared with other activities in the course	1	1	1	0	0.018	Reject H_{0A}

Table 6: Mann-Whitney test Results to the perceived learning

AQ2: Is there a difference in students' experience between the experimental and the control group?

The students' experience using the games/exercises was evaluated in terms of usability, confidence, challenge, satisfaction, social interaction, fun, focused attention, and relevance, based on the MEEGA+ model [Petri, 18a] (Table 7). Analysing usability, the degree to which a product (games/exercises) can be used by specified users (students) to achieve specified goals with effectiveness and efficiency [Petri, 18a], we can observe that, in general, both groups of students indicated that the games and the exercises used in the treatments provide a good usability (M_2), with a slightly higher usability indicated by the students from the experimental group in terms of learnability and easiness to play the games. Yet, although there is no difference between the groups in terms of aesthetics and accessibility, most of the measurement items present a difference between the groups, indicating that the usability of the games was perceived as better than the usability of the exercises. Therefore, we can reject the null hypothesis H_{0B}, confirming that the students in the experimental group perceived a better usability in the games than the control group.

In terms of confidence (M_3), evaluating if the students are able to make progress in the study of educational content through their effort and ability [Petri, 18a], the results of testing the hypothesis H_{0C} indicate that there is no difference between the experimental and control group. Thus, the null hypothesis H_{0C} is not rejected. This indicates that students of both groups felt confident that they were learning when playing the games as well as doing the exercises. On the other hand, analysing the challenge (M_4) provided by the games and the exercises with respect to the learner's competency level [Petri, 18a], results confirm the rejection of the null hypothesis H_{0D}. Thus, indicating that the students of the experimental group perceived this issue more positively than the students of the control group. This result also confirms other studies indicating that challenge is a main characteristic of educational games, which directly contributes to the students' learning [Hamari, 16].

In terms of satisfaction (M_5), evaluating if students feel that the dedicated effort in the games/exercises results in learning [Petri, 18a], the results of testing the hypothesis H_{0E} indicate that there is a significative difference in terms of satisfaction between the groups. Therefore, we can reject the null hypothesis H_{0E}. However, considering the recommendation of the games/exercises, a statistical difference could not be identified.

Regarding the social interaction (M_6), evaluating the feeling of a shared environment and interaction with other people [Petri, 18a], the results of testing the H_{0F} indicate that there is a statistical difference between the groups. Therefore, we can reject the null hypothesis H_{0F}, confirming that the games provide a better social interaction compared to the exercises. Similarly, in terms of fun (M_7), evaluating the pleasure, happiness, relaxing and distraction of the students during the activities [Petri, 18a], the results of testing the hypothesis H_{0G}, indicate that the students that played the games had a better perception of fun than the students of the control group, thus rejecting the null hypothesis H_{0G}. These results are similar to the findings of related studies, confirming that these features in games contribute to the creation of an enjoyable learning environment resulting in engagement and immersion, thus, contributing to learning [Prensky, 07; Hamari, 16; Cagiltay, 15].

Regarding the focused attention (M_8), which evaluates the concentration and the temporal dissociation of the students, and relevance, which evaluates if students realize that the educational proposal is consistent with their goals, the results of testing the hypothesis H_{0I} are indicating that the students of the experimental group had a better perception of focused attention and relevance than the students of the control group, thus, rejecting the null hypotheses H_{0H} and the H_{0I}.

In summary, we can identify a statistical difference of the students' experience between the experimental and control group. Students of the experimental group had a better perception of challenge, usability, satisfaction, social interaction, fun, focused attention, and relevance playing the games. Therefore, educational games may be considered an adequate instructional strategy for SPM education, contributing positively to the students' experience.

Questionnaire item	Control Group		Experimental Group		Mann-Whitney test	
	Median	Interquartile Range	Median	Interquartile Range	p-value (2-tailed)	Test result
Usability (M_2)						
The games/exercises design is attractive.	1	1	1	1	0.015	Reject H_{0B}
The text font and colours of the games/exercises are well blended and consistent.	1	2	1	1	0.621	Not Reject H_{0B}
I needed to learn a few things before I could play the games/start doing the exercises.	1	1	1	1	0.000	Reject H_{0B}
Learning to play these games/to do these exercises was easy for me.	1	1	2	1	0.000	Reject H_{0B}
I think that most people would learn to play these games/to do these exercises very quickly.	1	2	1	1	0.023	Reject H_{0B}
I think that these games/exercises are easy to play.	1	1	2	1	0.000	Reject H_{0B}
The game rules/explanations for doing the exercises are clear and easy to understand.	1	1	1	1	0.350	Not Reject H_{0B}
The fonts (size and style) used in the games/exercises are easy to read.	1	1	1	1	0.993	Not Reject H_{0B}
The colours used in the games/exercises are meaningful.	0	1	1	2	0.151	Not Reject H_{0B}
Confidence (M_3)						
When I first looked at the games/exercises, I had the impression that it would be easy for me.	1	1	2	2	0.127	Not Reject H_{0C}
The contents and structure helped me to become confident that I would learn with these games/exercises.	1	1	1	1	0.099	Not Reject H_{0C}
Challenge (M_4)						
These games/exercises are appropriately challenging for me.	0	2	1	0	0.000	Reject H_{0D}
These games/exercises provide new challenges (offers new obstacles, situations or variations) at an appropriate pace.	1	1	1	2	0.001	Reject H_{0D}
These games/exercises do not become monotonous as it progresses (repetitive or boring tasks).	1	2	1	2	0.000	Reject H_{0D}
Satisfaction (M_5)						
Completing the games tasks/exercises gave me a satisfying feeling of accomplishment.	1	1	1	1	0.000	Reject H_{0E}
It is due to my personal effort that I managed to advance in the games/exercises.	1	1	1	1	0.001	Reject H_{0E}
I feel satisfied with the things that I learned from the games/exercises.	1	1	1	0	0.018	Reject H_{0E}
I would recommend these games/exercises to my colleagues.	1	1	1	1	0.494	Not Reject H_{0E}
Social Interaction (M_6)						
I was able to interact with other people during the games/exercises.	0	1	2	1	0.000	Reject H_{0F}
The games/exercises promote cooperation and/or competition among the colleagues.	0	1	2	1	0.000	Reject H_{0F}
I felt good interacting with other people during the games/exercises.	0	1	2	1	0.000	Reject H_{0F}
Fun (M_7)						
I had fun with the games/exercises.	0	1	1	1	0.000	Reject H_{0G}
Something happened during the games/exercises which made me smile.	0	1	1	1	0.000	Reject H_{0G}
Focused attention (M_8)						
There was something interesting at the beginning of the games/exercises that captured my attention.	0	1	1	1	0.004	Reject H_{0H}
I was so involved in the gaming tasks/doing the exercises that I lost track of time.	-1	1	2	2	0.000	Reject H_{0H}
I forgot about my immediate surroundings while playing these games/doing the exercises.	0	1	0	1	0.000	Reject H_{0H}
Relevance (M_9)						
The games/exercises contents are relevant to my interests.	1	1	1	1	0.041	Reject H_{0I}
It is clear to me how the contents of the games/exercises are related to the course.	1	1	2	1	0.011	Reject H_{0I}
These games/exercises are an adequate teaching method for this course.	1	1	1	1	0.014	Reject H_{0I}
I prefer learning with these games/exercises to learning through other ways (e.g. other teaching methods).	1	1	1	2	0.001	Reject H_{0I}

Table 7: Mann-Whitney test Results to students' experience

AQ3: Is there a difference in students' learning between the experimental and the control group?

The two-way ANOVA was adopted to analyse the influence of the pre/post-test scores on learning effectiveness and also to assess if there is any interaction between these variables (test scores and learning) [Trochim, 08]. Based on the results (Table 8), we can observe that there is no interaction between the variables, thus, the factors (groups and period/time) may be analysed separately.

Comparing the experimental and control group, there is no statistical difference (p=0.122) and comparing the period/time (pre/post), there is a difference, representing the effect of the time (p<0.001). Thus, the results indicate that there is no statistical difference on learning effectiveness between the experimental and control group. As also shown in Figure 8, the means of tests scores (pre/post) of both groups are increasing over the time (period). Therefore, a statistical difference in learning effectiveness between the control and experimental group could not be identified.

	F	p-value
Group (Experimental/Control)	2.491	0.122
Period/Time (Pre/Post)	48.549	<0.001
Group*Period/Time	0.002	0.963

Table 8: Two-way analysis of variance

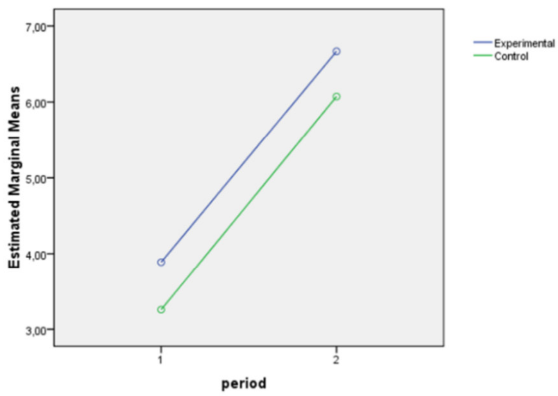


Figure 8: Estimated marginal means of experimental and control group

In addition, in order to explore this analysis, we also used the Fisher's Least Significant Difference (LSD) test [Trochim, 08], which compares the mean of one group with the mean of another (Table 9).

Group	**Test**	
	Pre	**Post**
Experimental	3.889 Aa	6.667 Ab
Control	3.259 Aa	6.074 Ab

Table 9: Fisher's LSD test results

Analysing Table 9, we can observe that the means of both groups increase in the post-test. In the pre-test, the Fisher's LSD test results (same letters (Aa)) indicate that there is no statistical difference between the groups. In the post-test, the Fisher's LSD test results (different letters (Ab)), indicate that there is statistical difference, confirming the time effect in the post-test scores in both, control and experimental group. Therefore, regarding the learning effectiveness, the null hypothesis H_{0J} is not rejected, indicating that, in this study there is no statistical difference with respect to the use of games and exercises teaching SPM. Both strategies (games/exercises) may be considered instructional strategies that contribute to the learning effectiveness in SPM education.

AQ4: Is there a difference in students' experience between different game genres?

We analysed the effect in students' experience using different game genres. Students' experience is evaluated in terms of usability, confidence, challenge, satisfaction, social interaction, fun, focused attention, and relevance. In order to analyse this question, we are considering only results from the experimental group, in order to identify difference between the game genres. We used the Mann-Whitney test, to analyse data from different game genres (quiz and simulation) and with a sample size variation [Trochim, 08]. The analysis results for each metric are presented in Table 10.

In general, we can observe that most of the test results are not rejecting the hypotheses H_{0L} to H_{0R}. These results indicate that based on the data no a statistical difference between quiz games and simulation games is identified in terms of confidence (M_{12}), challenge (M_{13}), satisfaction (M_{14}), social interaction (M_{15}), fun (M_{16}), focused attention (M_{17}), and relevance (M_{18}). Therefore, based on our results, both game genres may be considered adequate strategies to be included in SPM education, contributing to the students' experience.

However, in terms of usability (M_{11}), a statistical difference has been observed, confirming the rejection of the null hypothesis H_{0K}. This indicates that, although both game genres provide good usability, a slightly higher usability has been perceived by the students playing quiz games. Thus, in terms of usability the game genre may have an influence. The simulation game SCRUMIA is a non-digital game that requires the use of materials such as task board, user stories, pen, tape, and recycled paper, etc. in order to plan and execute a sprint of a hypothetical project applying SCRUM. On the other hand, one of the quiz games adopted in the study is a digital game developed based on the Kahoot! platform where players have to answer each question using their smartphone. Therefore, this result may also be have been influenced by the intrinsic characteristics of digital games, which provide rich interfaces to the users in terms of usability, confirming the rejection of the null hypothesis H_{0K}. Therefore, in order to identify more clearly the underlying factors that influence usability (game genre or game platform), further studies need to be conducted with a larger and more varied set of different games.

	Quiz Game		Simulation Game		Mann-Whitney test	
Questionnaire item	**Median**	**Interquartile Range**	**Median**	**Interquartile Range**	**p-value (2-tailed)**	**Test result**
Usability (M_{11})						
The games/exercises design is attractive.	1	1	1	2	0.003	Reject H_{0K}
The text font and colours of the games/exercises are well blended and consistent.	1	2	0	2	0.021	Not Reject H_{0K}
I needed to learn a few things before I could play the games/start doing the exercises.	2	1	1	1	0.003	Reject H_{0K}
Learning to play these games/to do these exercises was easy for me.	2	0	1	1	0.000	Reject H_{0K}
I think that most people would learn to play these games/to do these exercises very quickly.	2	1	1	1	0.000	Reject H_{0K}
I think that these games/exercises are easy to play.	2	1	1	2	0.001	Reject H_{0K}
The game rules/explanations for doing the exercises are clear and easy to understand.	2	1	1	1	0.008	Not Reject H_{0K}
The fonts (size and style) used in the games/exercises are easy to read.	2	1	1	1	0.001	Reject H_{0K}
The colours used in the games/exercises are meaningful.	1	2	0	1	0.002	Reject H_{0K}
Confidence (M_{12})						
When I first looked at the games/exercises, I had the impression that it would be easy for me.	1	1	1	1	0.073	Not Reject H_{0L}
The contents and structure helped me to become confident that I would learn with these games/exercises.	1	1	1	2	0.870	Not Reject H_{0L}
Challenge (M_{13})						
These games/exercises are appropriately challenging for me.	1	1	1	1	0.717	Not Reject H_{0M}
These games/exercises provide new challenges (offers new obstacles, situations or variations) at an appropriate pace.	1	1	1.5	2	0.126	Not Reject H_{0M}
These games/exercises do not become monotonous as it progresses (repetitive or boring tasks).	1	2	1.5	1	0.070	Not Reject H_{0M}
Satisfaction (M_{14})						
Completing the games tasks/exercises gave me a satisfying feeling of accomplishment.	1	0	1	1	0.456	Not Reject H_{0N}
It is due to my personal effort that I managed to advance in the games/exercises.	1	1	1	2	0.677	Not Reject H_{0N}
I feel satisfied with the things that I learned from the games/exercises.	1	0	1	1	0.341	Not Reject H_{0N}
I would recommend these games/exercises to my colleagues.	1	0	1	1	0.642	Not Reject H_{0N}
Social Interaction (M_{15})						
I was able to interact with other people during the games/exercises.	2	1	2	1	0.145	Not Reject H_{0O}
The games/exercises promote cooperation and/or competition among the colleagues.	2	1	2	1	1.000	Not Reject H_{0O}
I felt good interacting with other people during the games/exercises.	2	1	1.5	1	0.791	Not Reject H_{0O}
Fun (M_{16})						
I had fun with the games/exercises.	1	1	1.5	1	0.742	Not Reject H_{0P}
Something happened during the games/exercises which made me smile.	1	1	1	1	0.733	Not Reject H_{0P}
Focused attention (M_{17})						
There was something interesting at the beginning of the games/exercises that captured my attention.	1	1	1	1	0.699	Not Reject H_{0Q}
I was so involved in the gaming tasks/doing the exercises that I lost track of time.	1	2	1	1	0.640	Not Reject H_{0Q}
I forgot about my immediate surroundings while playing these games/doing the exercises.	0	2	1	1	0.183	Not Reject H_{0Q}
Relevance (M_{18})						
The games/exercises contents are relevant to my interests.	1	1	1	1	0.481	Not Reject H_{0R}
It is clear to me how the contents of the games/exercises are related to the course.	2	1	2	1	0.367	Not Reject H_{0R}
These games/exercises are an adequate teaching method for this course.	1	1	2	1	0.273	Not Reject H_{0R}
I prefer learning with these games/exercises to learning through other ways (e.g. other teaching methods).	1	2	1.5	1	0.115	Not Reject H_{0R}

Table 10: Mann-Whitney test results on students' experience using different game genres

AQ5: Is there a difference in students' learning between different game genres?

We adopted the two-way ANOVA to analyse the influence of the pre/post-test scores on learning effectiveness when using different games genres (quiz and simulation). Based on the results (Table 11), we can observe that there is interaction between the variables, thus, the factors (game genre and period/time) may be analysed together.

Analysing the results (Table 11), we can observe that there is statistical difference in terms of game genre (p=0.005) and period/time (pre/post) (p<0.001). Thus, these results are indicating that there is statistical difference on learning effectiveness depending on the game genres adopted in the study (quiz and simulation). As also shown in Figure 9, the means of tests scores (pre/post) of both game genres are increasing over time. However, a higher increase has been identified after playing the simulation games. Therefore, a statistical difference in learning effectiveness between the genre of quiz and simulation games can be identified.

	F	p-value
Game genre (Quiz/Simulation)	3.787	0.005
Period/Time (Pre/Post)	94.797	<0.001
Game genre*Period/Time	7.153	0.010

Table 11: Two-way analysis of variance

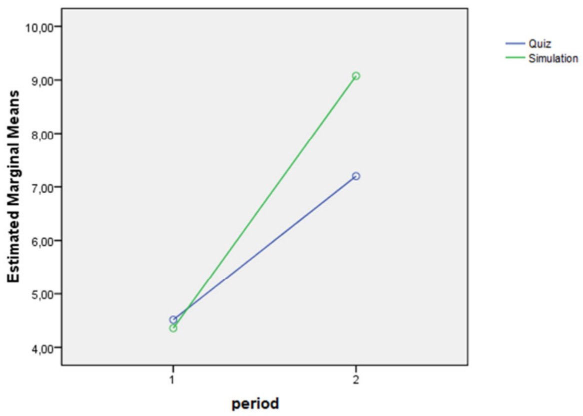


Figure 9: Estimated marginal means to quiz and simulation games

We also used the Fisher's Least Significant Difference (LSD) test in order to confirm our results, comparing the means of using quiz games with the means of using the simulation game (Table 12).

Game genre	**Test**	
	Pre	**Post**
Quiz	4.514 Aa	7.200 Ab
Simulation	4.357 Aa	9.077 Bb

Table 12: Fisher's LSD test results

Based on the results (Table 12), we can observe that the means with respect to both game genres increase in the post-test, yet with a higher increase in case of the simulation game, confirming our results presented in Table 11. In the pre-test, the Fisher's LSD test results (same letters (Aa)) indicate that there is no statistical difference between the game genres. However, in the post-test, the Fisher's LSD test results indicate that there is a statistical difference, confirming the time effect, similar in both game genres (same lower-case letters (b) and the increase in the performance in the post-test scores, mainly with respect to the simulation game (different upper-case letters (A and B)).

Therefore, in terms of learning effectiveness, the null hypothesis H_{0S} is rejected, indicating that, although both game genres contribute to the students' learning, the simulation game used in this study (SCRUMIA) provides a more positive effect on the students' learning when compared to the quiz games.

5 Discussion

The results of our study provide an initial indication that educational games may be an effective instructional strategy for SPM education, showing statistical evidence that games provide a positive contribution to the students' experience and perceived learning. However, based on the results, we cannot confirm that the games had a better learning effect when compared to exercises. Although the results show an increase in the students' performance (test scores) over the time, the same result was observed in the control group.

A possible explanation for this result may be the characteristics of the games used in the experiment. The games PMQuiz [Petri, 16] and PM Master [Gresse von Wangenheim, 12b] are quiz games, with a similar objective to review and reinforce basic SPM concepts. In this respect, the learning objective of both games is also similar to the objectives of the exercises (quizzes) answered by the students of the control group via Moodle. Although game elements like interaction and competitiveness are evidently different to the exercises, their similar objective and mechanics may be an explanation for the similar results in the students' performance. Another explanation may be the number of games played during the semester and the time spent in playing the games. Only three games were applied, and each game was played in one session of 30 minutes (PMQuiz) or 90 minutes (PM Master and SCRUMIA). The use of more games covering other SPM contents or learning levels and/or an application of the games with successive sessions may provide a better effect on students' learning. Other reasons may be related to the small sample size of the study, which may have not allowed an identification of a statistical difference between the groups, even adopting robust statistical methods to minimize this issue [Trochim, 08].

Analysing the results of the different genres of the games used in the study: quiz games (PMQuiz and PM Master) and a simulation game (SCRUMIA), we can observe that the simulation game provides a more positive result in terms of students' learning when compared to the quiz games, rejecting the null hypothesis H_{0S}. This result may be related to the intrinsic characteristics of the game genres. In simulation games, the students are typically in control of a certain environment or activity, which seeks to be as realistic as possible, while on the other hand, in quiz games, students just answer questions related to a particular knowledge [Djaouti, 11; Prensky, 07]. However, comparing the results of the simulation game (experimental group) to the control group, no statistical difference has been identified.

Similar results analysing the learning effectiveness in educational games were observed in related studies. The study evaluating X-MED [Gresse von Wangenheim, 09a], a computer-based educational game to practice the application of software measurement in the context of project management, was also not able to confirm a learning effect, possibly due to a small sample size and the small number of questions in the pre/post-test [Gresse von Wangenheim, 09a]. In a comprehensive evaluation of the SimSE game [Navarro, 07], a computer-based environment to simulate realistic game-based software process simulation models, a learning effect could also not be identified in the gaming group when compared to a lecture and a reading group. The lack of previous knowledge combined with the inadequate instructions given to the students on how to play SimSE were reported as the reasons for this result [Navarro, 07]. On the other hand, results of a pilot study evaluating SimSE indicate that the students felt that the game was effective at teaching software process concepts [Navarro, 07]. These results from different studies confirm that the evaluation of learning effectiveness is a complex task [Navarro, 07] due to several reasons as well as the difficult to isolate the factors that can influence the students' learning over a course may impact on the statistical results [Navarro, 07], [Wohlin, 12]. As a result, further large-scale studies with different types of games of longer duration or repeated application are necessary to analyse this question.

On the other hand, based on the perception of the students, we identified statistical evidence that most students believe that playing the games helped them to learn. Analysing the perceived learning, results of testing the H_{0A} hypothesis, show that the students in the experimental group had a better perception of learning than the students of the control group. Similar results have also been reported in the related studies focusing on specific SPM knowledge [Navarro, 07; Gresse von Wangenheim, 13b; Gresse von Wangenheim, 12b; Petri, 18]. Thus, as result of our study, we may extend this by providing evidence that games may contribute to students' perception of learning also over a SPM course. In general, our findings indicate that games used over a SPM course stimulate the students' motivation and active participation in the learning tasks, thus, contributing to the students' perception of learning.

Results of our study in terms of students' experience, testing the hypotheses H_{0B} to H_{0I}, show statistical evidence that games promote an engaged experience to the players over a SPM course, providing challenges, fun and creating an environment of cooperation and social interaction. These findings confirm similar results of related studies [Gresse von Wangenheim, 13a; Gresse von Wangenheim, 14; Gresse von

Wangenheim, 13b; Gresse von Wangenheim, 12a] when evaluating a game to teach a specific SPM knowledge. Therefore, in the context of our study, comparing games to traditional exercises over a SPM course, we can confirm that games provide a better experience for the students than the traditional exercises in the context of SPM education.

The findings of our study related to the specific context of SPM education, in general, are similar, to results reported with respect to educational/serious games in different knowledge areas, such as health, emergency training, science, etc. [Virvou, 05; Connolly, 12; Capuano, 15; AL-Smadi, 18]. Mainly related to the students' experience provided by the games, it is possible to confirm that games contribute to the creation of an enjoyable learning environment resulting in engagement and immersion, contributing to the students' learning.

5.1 Threats to validity

Due to the characteristics of this study, it is subject to some threats to validity, including threats to internal, external, construct and conclusion validity [Wohlin, 12]. We, therefore, identified potential risks/threats and applied mitigation strategies in order to minimize their impact on our research.

Internal validity. A threat refers to the groups being composed of students from different undergraduate programs with different backgrounds. However, this threat is minimized due to the fact that students of both groups are similar in terms of gender, age group and computing background. In addition, both SPM courses are similar in terms of objectives, duration, SPM contents, and year that they are offered in the programs. Another threat may be the variation of the sample size of the experimental and control group. Therefore, adequate statistical methods were adopted considering the difference of the sample sizes. Another risk may be the difference in terms of difficulty in the pre/post-tests. In this respect, both pre/post-tests were carefully designed and reviewed by all authors together creating tests with similar content and difficulty.

External validity. A threat to the possibility to generalize the results is related to the sample size. Although this study was conducted with a small sample size, this number reflects the small number of students commonly enrolled in computing courses [Bowman, 18]. However, considering our sample size, we choose adequate statistical methods in order to correctly analyse the experiment results with statistical significance. Another threat refers to the extent to which the experiment is dependent on the specific researchers. Although the experiment was run under supervision of a group of researchers from the same research group, we tried to minimize this threat by defining and following a systematic research method, adopting the GQM approach to clearly define the study objective, the process of data collection, and data analysis.

Construct validity. Another threat refers to the degree to which the measures are accurately measuring the concepts that they purport to measure. In order to minimize this threat, we adopted the MEEGA+ evaluation model [Petri, 18a], a systematic, reliable and valid model for the evaluation of games, which evaluates games in terms of player's experience and perceived learning based on the students' perception after they played a game through a standardized questionnaire. The MEEGA+ questionnaire was adopted to collect data from the students of the experimental group and an adaptation of the MEEGA+ questionnaire was used to collect data from the students of the control group.

Conclusion validity. Threats to the conclusion validity were minimized by adopting the MEEGA+ questionnaire, which provides statistical evidence that it is, in fact, measuring what it proposes to measure [Petri, 18a]. In addition, in order to analyse the data collected by the MEEGA+ questionnaire and the pre/post-test, adequate statistical methods were chosen based on the characteristics of our experiment. The choice of the statistical methods and the interpretation of the study results were revised by a senior researcher with expertise in statistics.

6 Conclusions

This article presents an experimental study to analyse the effectiveness of educational games in students' experience and learning in SPM courses of higher computing education. Although the results do not provide sufficient evidence for learning effectiveness based on a comparison of pre/post-test results of both groups, we were able to demonstrate that educational games may be considered an effective instructional strategy for SPM education, based on the perceived effect on learning and students' experience. By playing the games, students in the experimental group had a better perception of learning than the students of the control group, answering the exercises. This perception may have been potentiated by the students' experience provided by the games. Analysing the students' experience of both groups, the results show statistical evidence that students in the experimental group felt challenged by the games, keeping their attention focused in the learning activities, providing an environment of cooperation and competition among the students, and contributing to their engagement and immersion in the learning tasks provided by the games.

Therefore, we can conclude that, although, the study could not statistically demonstrate a learning effect, games may be considered an effective instructional strategy for SPM education, providing a positive contribution to the students' perceived learning, as well as providing an engaging experience for computing students. To further extend this research, we are planning repeat the experiment with modifications in the treatments in order to obtain a better understanding of the learning effectiveness. We also intend to include a greater variety of games enabling the analysis of differences with respect to different game genres. In addition, we are also planning to analyse the correlation between measures of perceived learning and the learning effectiveness when adopting games for SPM/computing education.

Acknowledgements

This work was supported by the CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico – www.cnpq.br), an entity of the Brazilian government focused on scientific and technological development.

References

[ACM, 13] ACM/IEEE-CS.: Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science, 2013. Available on: <https://www.acm.org/education/CS2013-final-report.pdf> Access: 06 Feb. 2019.

[AL-Smadi, 18] AL-Smadi, M., Capuano, N., Guetl, C.: Decoupling assessment and serious games to support guided exploratory learning in smart education. Journal of Ambient Intelligence and Humanized Computing, 9 (3), 497-511, 2018.

[All, 16] All, A., Castellar, E. P. N., Looy, J. V.: Assessing the effectiveness of digital game-based learning: Best practices, Computers & Education, 92–93, 90-103, 2016.

[Backlund, 13] Backlund, P., Hendrix, M.: Educational games - Are they worth the effort? A literature survey of the effectiveness of serious games, Proc. of the 5th Int. Conf. on Games and Virtual Worlds for Serious Applications, Poole, GB, 2013.

[Basili, 94] Basili, V. R., Caldiera, G., Rombach, H. D.: Goal, Question Metric Paradigm. In J. J. Marciniak, Encyclopedia of Software Engineering, pp. 528-532, Wiley-Interscience, New York, NY, USA, 1994.

[Battistella, 16] Battistella, P., Wangenheim, C. G.: Games for Teaching Computing in Higher Education – A Systematic Review. IEEE Technology and Engineering Education (ITEE) Journal, 9(1), 8-30, 2016.

[Bavota, 12] Bavota, G., Lucia, A., Fasano, F., Oliveto, R., Zottoli, C.: Teaching software engineering and software project management: an integrated and practical approach. Proc. of the International Conference on Software Engineering, pp. 1155-1164, Piscataway, USA, 2012.

[Black, 09] Black, R. T., Osipov, I. A.: Gaming in the education of project managers. Proc. PMI® Global Congress, Orlando, USA, 2009.

[Bourque, 14] Bourque, P., Fairley, R. E.: Swebok v3.0 Guide to the software engineering body of knowledge. IEEE Computer Society, 2014.

[Bowman, 18] Bowman, D. D.: Declining Talent in Computer Related Careers. Journal of Academic Administration in Higher Education, 14(1), 1-4, 2018.

[Cagiltay, 15] Cagiltay, N. E., Ozcelik, E., Ozcelik, N. S.: The effect of competition on learning in games. Computers & Education, 87, 35-41, 2015.

[Calderón, 13] Calderón, A., Ruiz, M.: ProDec: a serious game for software project management training. Proc. of the 8th International Conference on Software Engineering Advances, pp. 565-570. Venice, Italy, 2013

[Calderón, 15] Calderón A., Ruiz M.: A systematic literature review on serious games evaluation: An application to software project management, Computers & Education, 87, 396-422, 2015.

[Calderón, 17] Calderón A., Ruiz M., Orta E.: Integrating Serious Games as Learning Resources in a Software Project Management Course: The Case of ProDec, Proc. of the 1st International Workshop on Software Engineering Curricula for Millennials, pp. 21-27, Buenos Aires, Argentina, 2017.

[Calderón, 18a] Calderon, A., Trinidad, M., Ruiz, M., O'Connor, R. V.: Towards a Standard to Describe and Classify Serious Games as Learning Resources for Software Project Management. In: Larrucea X., Santamaria I., O'Connor R., Messnarz R. (Eds) Systems, Software and Services Process Improvement. Communications in Computer and Information Science, vol 896. Springer, Cham, 2018.

[Calderón, 18b] Calderón A., Ruiz M., O'Connor R. V.: A multivocal literature review on serious games for software process standards education, Computer Standards & Interfaces, 57, 36-48, 2018.

[Calderón, 19] Calderón A., Ruiz M., O'Connor R. V.: Designing game scenarios for software project management education and assessment, IET Software, 13(2), 144-151, 2019.

[Capuano, 15] Capuano N., King, R. Knowledge-based assessment in serious games: an experience on emergency training. Journal of e-Learning and Knowledge Society, 11(3), 2015.

[Connolly, 12] Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., Boyle, J. M.: A systematic literature review of empirical evidence on computer games and serious games, Computers & Education, 59(2), 661-686, 2012.

[Dantas, 04] Dantas, A. R., Barros, M. O. Werner, C. M. L.: A Simulation-Based Game for Project Management Experiential Learning. Proc. of the International Conference on Software Engineering and Knowledge Engineering, Alberta, Canada, 2004.

[Djaouti, 11] Djaouti, D., Alvarez, J., Jessel, J. P., Rampnoux, O.: Origins of Serious Games. In Ma, M., Oikonomou, A., Jain, L. (Eds). Serious Games and Edutainment Applications. London: Springer, 2011.

[Fernandes, 10] Fernandes, J. M., Sousa, S. M.: PlayScrum - A Card Game to Learn the Scrum Agile Method. Proc. of the International Conference on Virtual Worlds and Games for Serious Applications, Braga, Portugal, 2010, pp. 52-59.

[Geist, 07] Geist, D.B., Myers, M.E.: Pedagogy and project management: Should you practice what you preach. Journal of Computing Sciences in Colleges, 23(2), 202–208, 2007.

[Gresse von Wangenheim, 09a] Gresse von Wangenheim, C., Thiry, M. Kochanski, D.: Empirical evaluation of an educational game on software measurement. Empirical Software Engineering, 14(4), 418-452, 2009.

[Gresse von Wangenheim, 09b] Gresse von Wangenheim, C., Shull, F.: To Game or Not to Game? Software, IEEE, 26(2), 92-94, 2009.

[Gresse von Wangenheim, 12a] Gresse von Wangenheim, C.: PM Master. Available at: <http://www.gqs.ufsc.br/pm-master/> Access: 06 February 2018.

[Gresse von Wangenheim, 12b] Gresse von Wangenheim, C., Savi, R., Borgatto, A. F.: DELIVER! An Educational Game for Teaching Earned Value Management in Computing Courses. Information and Software Technology, 54(3), pp. 286-298, 2012.

[Gresse von Wangenheim, 13a] Gresse von Wangenheim, C., Carvalho, C O. P., Battistela, P. E.: Teaching Team Management in Software Project Management Courses. Brazilian Journal of Computers in Education, 21(1), 2013.

[Gresse von Wangenheim, 13b] Gresse von Wangenheim, C., Savi, R., Borgatto, A. F.: SCRUMIA - An educational game for teaching SCRUM in computing courses. Journal of Systems and Software, 86(10), 2675-2687, 2013.

[Gresse von Wangenheim, 14] Gresse von Wangenheim, C., Rausis, B., Soares, G., Savi, R., Borgatto, A. F.: Project Detective a Game for Teaching Earned Value Management. International Journal of Teaching and Case Studies, 5(3/4), 216-234, 2014.

[Hamari, 16] Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., Edwards, T.: Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. Computers in Human Behavior, 54, 170-179, 2016.

[Hussein, 15] Hussein, B. A.: A Blended Learning Approach to Teaching Project Management: A Model for Active Participation and Involvement: Insights from Norway. Education Sciences, (5), 104-125, 2015.

[Kosa, 16] Kosa, M., Yilmaz, M., O'Connor, R., Clarke, P.: Software engineering education and games: a systematic literature review. Journal of Universal Computer Science, 22(12), 1558-1574, 2016.

[Navarro, 07] Navarro, E. O., van der Hoek, A.: Comprehensive evaluation of an educational software engineering simulation environment. Proc. of the 20th Conference on Software Engineering Education & Training, pp.195-202. Dublin, Ireland, 2007.

[Ojiako, 11] Ojiako, U., Ashleigh, M., Chipulu, M., Maguire, S.: Learning and teaching challenges in project management. International Journal of Project Management, (29), 268–278, 2011.

[Petri, 16] Petri, G., Battistella, P. E., Cassettari, F., Gresse von Wangenheim, F., Hauck, J. C. R.: A Quiz Game for Knowledge Review on Project Management. Proc. of the 27th Brazilian Symposium on Informatics in Education. Uberlândia, MG, Brazil (in Portuguese), 2016.

[Petri, 17a] Petri, G., Gresse von Wangenheim, C.: How games for computing education are evaluated? A systematic literature review. Computers & Education, 107, 68-90, 2017.

[Petri, 17b] Petri, G., Gresse von Wangenheim, C., Borgatto, A. F.: Quality of Games for Teaching Software Engineering: An Analysis of Empirical Evidence of Digital and Non-digital Games. Proc. of the International Conference on Software Engineering, pp. 150-159, Buenos Aires, Argentina, 2017.

[Petri, 18a] Petri, G., Gresse von Wangenheim, C., Borgatto, A. F.: MEEGA+, Systematic Model to Evaluate Educational Games. In Newton Lee (Eds) Encyclopedia of Computer Graphics and Games. Springer, Cham, pp. 1-7, 2018.

[Petri, 18b] Petri, G., Calderón, A., Gresse von Wangenheim, C., Borgatto, A. F., Ruiz, M.: Games for Teaching Software Project Management: An Analysis of the Benefits of Digital and Non-Digital Games. Journal of Universal Computer Science, 24(10), 1424-1451, 2018.

[Pfahl, 01] Pfahl, D., Ruhe, G., Koval, N.: An Experiment for Evaluating the Effectiveness of Using a System Dynamics Simulation Model in Software Project Management Education, Proc. of the 7th Int. Symposium on Software Metrics, pp.97-109, London, GB, 2001.

[PMI, 13] PMI – Project Management Institute, A Guide to the Project Management Body of Knowledge, 5. ed., Newtown Square, 2013.

[Prensky, 07] Prensky, M.: Digital Game-Based Learning. New York: Paragon House, 2007.

[Randel, 92] Randel J. M., Morris B. A. Wetzel, C. D. Whitehill B. V.: The Effectiveness of Games for Educational Purposes: A Review of Recent Research. Simulation & Gaming, 23(3), 261–276, 1992.

[Ritterfeld, 10] Ritterfeld, U., Cody, M., Vorderer, P. (Eds.): Serious Games. New York: Routledge, 2010.

[Rumeser, 18] Rumeser, D., Emsley, M.: Project Management Serious Games and Simulation: A comparison of three learning methods. The Journal of Modern Project Management, 5(3), 2018.

[Trochim, 08] Trochim, W. M., Donnelly, J. P.: Research methods knowledge base (3rd ed.). Mason, OH: Atomic Dog Publishing, 2008.

[Virvou, 05] Virvou M., Katsionis G., Manos K.: Combining Software Games with Education: Evaluation of its Educational Effectiveness. Educational Technology & Society, 8 (2), 54-65, 2005.

[Wohlin, 12] Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B., Wesslén, A.: Experimentation in Software Engineering, Springer-Verlag Berlin Heidelberg, 2012.