INCoD
Instituto Nacional para Convergência Digital

UNIVERSIDADE FEDERAL
DE SANTA CATARINA

ISSN 2236-5281

Technical Report
INCoD/GQS.05.2018.E

MEEGA+: A Method for the Evaluation of Educational Games for Computing Education

Authors:
Giani Petri
Christiane Gresse von Wangenheim
Adriano Ferreti Borgatto

Version 0.9
Status: Final
Distribution: External
JULY 2018

Brazilian Institute for Digital Convergence

Federal University of Santa Catarina - UFSC
Campus Universitário João David Ferreira Lima - Trindade
Departamento de Informática e Estatística - Room 320
Florianópolis-SC - CEP 88040-970

Fone / FAX: +55 48 3721-9516 R.17
www.incod.ufsc.br

ISSN 2236-5281

Relatório Técnico do Instituto Nacional para Convergência Digital/ Departamento de Informática e Estatística, Centro Tecnológico, Universidade Federal de Santa Catarina. -- v.1, n.1 (2011).-- Florianópolis: INE, UFSC, 2011 -

Semestral
Resumo em português

ISSN 2236-5281

1. Convergência digital. 2. Tecnologia da informação. 3. Informática na saúde. 4. Mídia digital I. Universidade Federal de Santa Catarina. Departamento de Informática e Estatística.

Abstract

Educational games are assumed to be an effective and efficient instructional strategy for computing education. However, it is essential to systematically evaluate such games in order to obtain sound evidence on their quality. A prominent evaluation model is MEEGA (Model for the Evaluation of Educational Games) providing a support to evaluate games in terms of motivation, user experience and learning. However, analyses of the initial version of the MEEGA model have identified limitations regarding its validity and reliability. Furthermore, a more comprehensive support is required in order to guide instructors and/or researchers in how to conduct game evaluations in order to obtain reliable and valid results. Thus, the objective of this technical report is to present the MEEGA+ method, which aims to evaluate the quality of educational games used as instructional strategy for computing education, improving the initial version of the MEEGA model. The MEEGA+ method has been developed taking also into consideration results from a systematic literature review on the evaluation of educational games. The MEEGA+ method is composed by the quality model, defining quality factors to be evaluated through a standardized measurement instrument, and defines the process to evaluate the quality of educational games used for computing education using the MEEGA+ model. The MEEGA+ method provides game creators, instructors and researchers with a systematic support in order to evaluate the quality of educational games and, thus, contribute to their improvement and effective and efficient adoption in practice.

1. Introduction

Computing has an important role in our society nowadays, strongly influencing the development of education, science, engineering, business, and many other knowledge areas. In this context, computing professionals will continue to play an essential role, changing and shaping our future (ACM/IEEE-CS, 2013). Therefore, it is important that computing education attracts quality students in order to prepare them to be qualified and responsible computing professionals for them to contribute to the development of diverse knowledge areas (Parsons, 2011; ACM/IEEE-CS, 2013).

Although computing systems/technologies and the profile of computing students have expressively changed in the past decades, most of the undergraduate computing courses are still taught in traditional ways (Parsons, 2011; Freeman et al., 2014). Yet, in today's world, traditional instructional strategies may not be adequate for effective and efficient learning (Parsons, 2011; Freeman et al., 2014; Dolan & Collins, 2015). For these reasons, more active instructional strategies are required, focusing on the students, allowing them to learn by doing and, thus, enabling more effective and efficient learning (Freeman et al., 2014; Dolan & Collins, 2015). In this context, games have been used as an instructional strategy in order to provide more practical learning opportunities to computing students (Kosa et al., 2016; Battistella & Gresse von Wangenheim, 2016; Kordaki & Gousiou, 2016; Calderón, Ruiz, & O'Connor, 2018). They are considered to be an adequate instructional strategy for teaching computing by contributing to students' learning, motivation, engagement, etc. (Boyle, Connolly, & Hainey, 2011; Connolly et al., 2012; Gibson & Bell, 2013; Kosa et al., 2016; Hamari et al., 2016; Calderón et al., 2018; Çiftci, 2018).

In order to confirm these expected benefits of games and obtain empirical evidence of games' quality, it is important to systematically evaluate such games. However, currently, there are only few approaches, which provide a systematic support for game evaluations (Calderón & Ruiz, 2015; All et al., 2016; Petri & Gresse von Wangenheim, 2016; Kordaki & Gousiou, 2017; Tahir & Wangmar, 2017; Santos, Souza, Figueiredo, & Dayrell, 2018). In this context, a prominent evaluation model is MEEGA (Model for the Evaluation of Educational Games) providing a support to evaluate games in terms of motivation, user experience and learning. However, analyses of the initial version of the MEEGA model, based on a sample of 723 responses, have identified limitations regarding its validity and reliability. These limitations are related to an overlap of theoretical concepts of the quality factors motivation and user experience (Petri et al., 2017). Thus, evaluations using the initial version of the MEEGA model may lead to imprecise results on the game's quality, not correctly identifying evidence of their benefits regarding the overlapped concepts (motivation and user experience). In addition, the MEEGA model does not provide a more comprehensive support, for example, defining a step by step in order to guide researchers in the planning, execution and analysis of results of game evaluations. Thus, evaluations using the initial version of the MEEGA model may be impaired, producing not valid results if the researcher does not have experience in the conduction of this kind of evaluation.

In this context, this technical report presents the MEEGA+ method, providing a comprehensive support for a systematic evaluation of games used for computing education. The MEEGA+ method is composed of an evaluation model (MEEGA+ Model) defining quality factors to be evaluated through a standardized measurement instrument, a scale, which classifies the evaluated game according to its quality level, and a process (MEEGA+ Process) defining phases, activities and work products, guiding researchers on how to plan, execute and analyse the results of game evaluations.

The MEEGA+ method is designed to be used by game developers, researchers, and instructors to evaluate the quality of games used for computing education. And, thus, contributing to

the analysis of which games provide an effective and efficient learning and, contributing positively to the learning of future computing professionals.
Although the emphasis of the MEEGA+ method is on the evaluation of games used for computing education, we assume that the MEEGA+ method can be used and adapted for the evaluation of games to teach others knowledge areas, as performed by Gomes (2016), Herpich et al. (2007), Pereira et al. (2017), Silva et al. (2017a) and Silva et al. (2017b). However, when transferring the method to other knowledge areas, further empirical studies are necessary to evaluate and confirm the reliability and validity of the MEEGA+ method also in these areas.

2. Research Method

In order to develop the MEEGA+ method, a multi-method research is adopted, as shown in Figure 1.

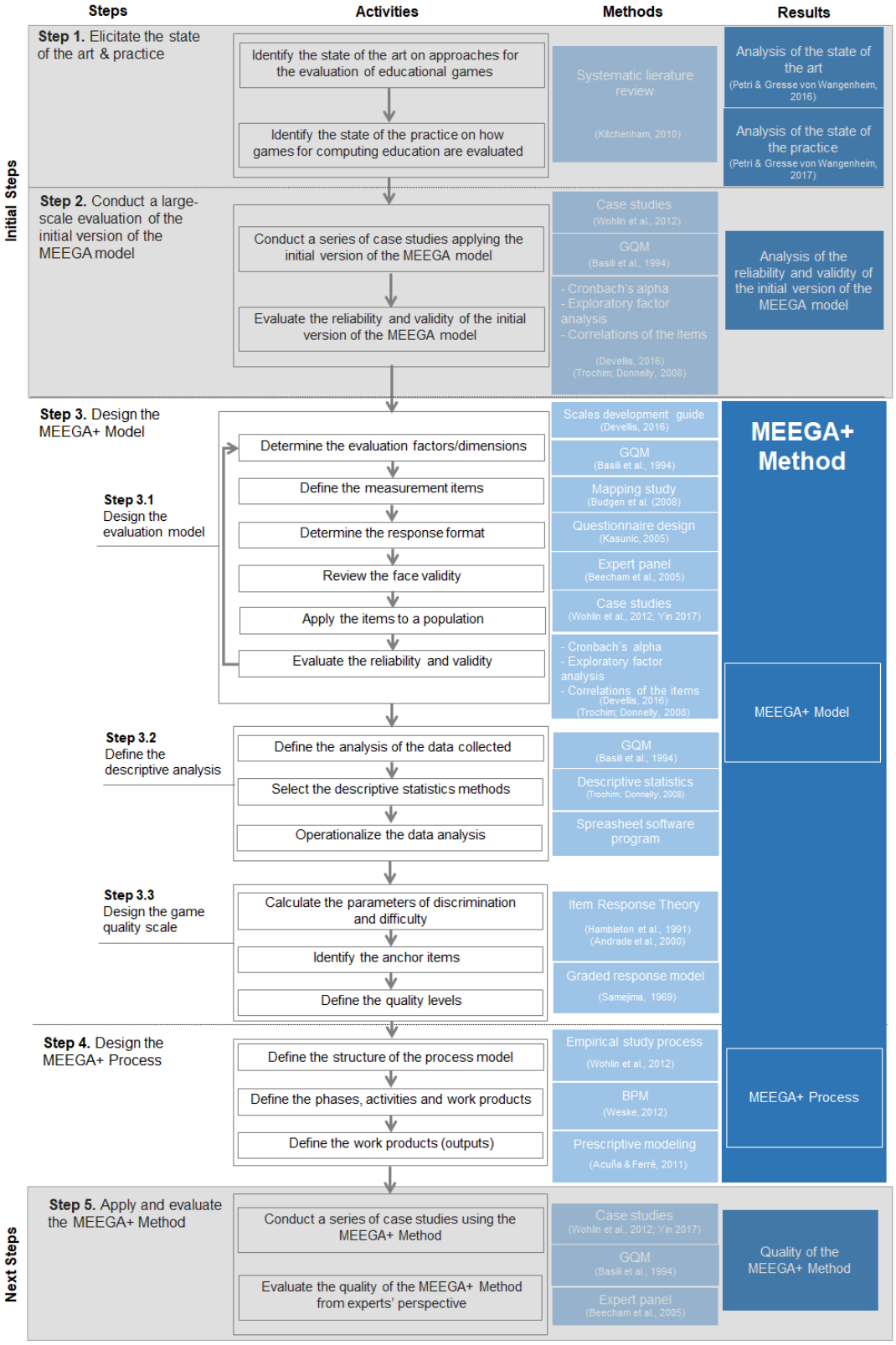

Figure 1. Research method

Preceding the design of the MEEGA+ method, we elicited the state of the art identifying existing approaches (methods, models, frameworks, and scales) for the systematic evaluation of educational games (Petri & Gresse von Wangenheim, 2016). We also analysed the state of practice, eliciting how games used for computing education are currently evaluated (Petri & Gresse von Wangenheim, 2017). As a result, we identified the MEEGA model (Savi et al., 2011) as a well-defined approach for the evaluation of games for computing education that is being widely used in practice.

We, then, conducted a large-scale study of the initial version of the MEEGA model analysing its validity and reliability. The analysis was conducted based on data collected in 43 case studies, evaluating 20 different games for computing education, involving a population of 723 students (Petri et al., 2017). As a result, we identified that the initial version of the MEEGA model is acceptable in terms of reliability (α=.915). However, in terms of its validity, a conceptual overlap with respect to the factors motivation and user experience has been observed, indicating a need for the redesign of the MEEGA model.

Detailed results and a complete description of these preceding steps can be found in Petri & Gresse von Wangenheim (2016) presenting the state of the art, Petri & Gresse von Wangenheim (2017) presenting the state of the practice, and Petri et al. (2017) presenting the large-scale study of the initial version of the MEEGA model.

In this technical report, we focus on present the results of Steps 3 (Design of the MEEGA+ Model) and 4 (Design of the MEEGA+ Model), composing as result the MEEGA+ Method.

Step 3. Design the MEEGA+ model.

Step 3.1 Design the evaluation model. The MEEGA+ model has been developed, as an evolution of the initial version of the MEEGA model (Savi et al., 2011), based on the results of the literature reviews (Petri & Gresse von Wangenheim, 2016; Petri & Gresse von Wangenheim, 2017) and the large-scale analysis of the initial version of the MEEGA model (Petri et al., 2017). The development of the MEEGA+ model follows the procedure of the scale development guide proposed by DeVellis (2016) and the guide for questionnaire design by Kasunic (2005):

Determine the evaluation factors/dimensions. Adopting the GQM (Goal/Question/Metric) approach (Basili et al., 1994), the evaluation objective is defined and systematically decomposed into factors to be measured. The factors are defined in order to support the development of the measurement instrument (questionnaire), based on a mapping study of their concepts following the procedure proposed by Budgen et al. (2008).

Define the measurement items. The measurement of the factors is operationalized by decomposing the factors into measurement instrument items. For the definition of the items we also took into consideration other standardized questionnaires found in literature. We analyzed the pool of items in terms of similarity and redundancy, customizing and unifying the selected items for the defined evaluation factors. In order to standardize the selected items, all items were refined and transformed into positive statements.

Determine the response format. In this step, the response format for the items of the measurement instrument is defined. This definition is based on response formats typically used in standardized questionnaires following the scale development guide proposed by DeVellis (2016).

Review the face validity. In this step, face validity (Trochim & Donnelly, 2008) is analysed by an expert panel (Beecham, Hall, Britton, Cotte, & Rainer, 2005). The expert panel is composed of a multidisciplinary group of senior researchers with backgrounds in computing and/or statistics. The review aims at analysing clarity, relevance, consistency, and completeness of the measurement instrument items of the MEEGA+ model. The suggestions of the experts, including changes in the wording and text formatting, were considered in the development of the MEEGA+ measurement instrument.

Apply the items to a population. In this phase, a series of case studies applying several games for computing education is conducted. In each case study, after the game session (treatment), the MEEGA+ measurement instrument is used for data collection on the participants' perceptions about the game. We use a non-probability sampling technique in each case study applying the convenience sampling method (Trochim and Donnelly, 2008), in which our sample is composed by students enrolled in computing courses.
Evaluate the reliability and validity. In the analysis phase, we pooled the data collected in each case study in a single sample for data analysis. Data were analyzed in terms of reliability and construct validity following the definition of Trochim and Donnelly (2008) and the scale development guide proposed by DeVellis (2016). In terms of reliability, we measured the internal consistency through the Cronbach's alpha coefficient (Cronbach, 1951). Construct validity was analyzed using an exploratory factor analysis and based on evidence of convergent and discriminant validity, obtained through the degree of correlations of the items (DeVellis, 2016; Trochim and Donnelly, 2008). The results of the statistical analysis were interpreted by researchers in the context of computing education in order to identify the reliability and validity of the MEEGA+ measurement instrument. Results of this step are used to identify the quality factors that represent the responses to the MEEGA+ measurement instrument.

Step 3.2 Define the descriptive analysis. In order to provide a support in the analysis of the data collected through the MEEGA+ measurement instrument, the analysis of the data collected is defined and analysed through descriptive statistics methods.
Define the analysis of the data collected. In accordance with the evaluation objective defined and the respective analysis questions, following the MEEGA+ model and the GQM approach (Basili et al., 1994), the analysis of the data collected is defined.
Select the descriptive statistics methods. Descriptive statistics methods are used to describe and graphically present interesting aspects of the data collected (Wohlin et al., 2012). Thus, in order to analyse data collected through the MEEGA+ measurement instrument, descriptive statistics methods are selected (Trochim & Donnelly, 2008; Wohlin et al., 2012), such as measures of central tendency (median, average and frequency of responses) and graphical visualization (frequency charts).
Operationalize the data analysis. In this step, the descriptive statistics methods selected are used to calculate the measures of the central tendency and provide a graphical visualization of the data collected. Thus, following the MEEGA+ model and the GQM approach, the descriptive statistics are interpreted in order to answer the defined analysis questions and, thus, achieving the evaluation objective.

Step 3.3 Design the game quality scale. In order to use the MEEGA+ measurement instrument to classify the quality level of the game being evaluated, a MEEGA+ scale is developed. The MEEGA+ scale aims to classify the evaluated game regarding to its quality level, based on students' perception, allowing to identify which requirements correspond to the lowest or highest level of quality. The MEEGA+ scale is developed by adopting the statistical technique Item Response Theory (IRT) (Pasquali & Primi, 2003), which allows to express through mathematical models the relationship between observable variables (questionnaire items) and latent traits (game's quality) based on the students' perceptions (Pasquali & Primi, 2003).
Calculate the parameters of discrimination and difficulty. In order to calculate the parameters of discrimination (*a*) and difficulty (*b*) we use the probabilistic model proposed by Samejima (1969). The model of Samejima (1969) is used due to the nature of the analysed data, distributed in categories on a gradual scale (Likert scale). The parameter *a* is associated with how much the item discriminates (differentiates) the students in relation to the latent trait (game's quality), where the higher its value, the more associated with the latent trait is the item. The parameter *b* is associated with the degree of difficulty of the item, where the

higher its value, the more difficult the students to agree with the item in relation to the game's quality.

Identify the anchor items. In order to identify the anchor items, which determine the categories of the latent trait (game's quality), we consider the probability parameter $Pi,k(\theta)>=0,50$ (Andrade et al., 2000).

Define the quality levels. In this step, the quality levels determined by the anchor items are defined considering the latent trait (game's quality), thus, defining and describing the quality levels of a measurement scale. The quality level of the scale is defined using an average 50 and standard deviation 15, scale (50.15), applying the formula $\theta_{50,15}=50+15*\theta_{0,1}$ (Andrade et al., 2000).

Step 4. Design the MEEGA+ process. The MEEGA+ process aims to provide a systematic support, guiding researchers and instructors, in the conduction of game evaluations.

Define the structure of the process model. The process is modeled in a prescriptive way (Acuña & Ferré, 2001). The goal of prescriptive modelling is to define the required or recommended means of executing the process, thus, defining how the process should be performed, establishing rules, guidelines and standards (Acuña & Ferré, 2001). The process is organized in phases, activities and work products (Acuña et al., 2000; Benali & Derniame, 1992; Finkelstein et al., 1994). A phase is a step of the process, presenting a set of activities in a structured sequence. Activities are the stages of a process that implements procedures to transform a product. Work products are the inputs and outputs of an activity from a process, they may be produced and consumed throughout the process. In order to provide a graphical notation of the MEEGA+ process, we adopt the Business Process Modeling Notation (BPMN) (Weske, 2012).

Define the phases, activities and work products. In this step, the phases, activities, and work products of the MEEGA+ process are defined, based on an empirical study process as proposed by Wohlin et al. (2012) and practical experiences in the conduction of game evaluations.

Define the work products (outputs). In this step, a template for each work product (output) generated during the process is defined. The templates are aimed to organize and standardize the information generated during the process in order to assist researchers in the definition, planning, execution, analysis, and presentation of the evaluation.

As next steps, we plan to conduct a series of case studies (Yin, 2009; Wohlin et al., 2012) evaluating games (digital and non-digital) for computing education using the MEEGA+ method. The case studies are planned to be conducted in different higher computing education courses, from various educational institutions in Brazil and abroad. In addition, we plan evaluate the quality of the MEEGA+ method from the expert's perspective. Thus, we plan evaluate the quality of the MEEGA+ method in terms of consistency, completeness, comprehensibility, comprehensibility, ambiguity, flexibility and usability (efficacy, efficiency, learnability, ease of use, and utility) (Siniscalco & Auriat, 2005; Davis, 1989; Rittgen, 2010; Matook; Indulska, 2009), through an expert panel (Beecham et al., 2005). By adopting the GQM approach (Basili et al., 1994), the evaluation is defined and decomposed into analysis questions and metrics, which are collected through a questionnaire answered by the experts after analysing the MEEGA+ method.

3. MEEGA+: A Method for the Evaluation of Games for Computing Education

This section presents the MEEGA+ method, which has been systematically developed for the quality evaluation of games used as an instructional strategy for computing education, based on the results of the literature reviews (Petri & Gresse von Wangenheim, 2016; Petri & Gresse von Wangenheim, 2017) and the large-scale analysis of the initial version of the MEEGA model (Petri et al., 2017). In this study, we define a method as a systematic approach to achieve a certain objective or result and, which describes the characteristics of an ordered process or a procedure used in the engineering of a product (IEEE, 2002; IEEE 2010). Based on this definition, the MEEGA+ method aims to provide a systematic support for the evaluation of games for computing education. It is composed of an evaluation model (MEEGA+ Model) defining quality factors to be evaluated through a standardized measurement instrument, and a scale, which classifies the evaluated game according to its quality level, and a process (MEEGA+ Process) defining phases, activities and work products, guiding researchers in how to plan, execute and analyse the results of game evaluations.
In this study, we consider that a quality model defines a set of characteristics and/or subcharacteristics and their relationships, providing a basis for the specification of quality requirements to be evaluated (ISO/IEC, 2011). A process is defined as a sequence of interrelated steps and activities that are performed to achieve a determined purpose (IEEE, 2002). Figure 2 presents the composition of the MEEGA+ method.

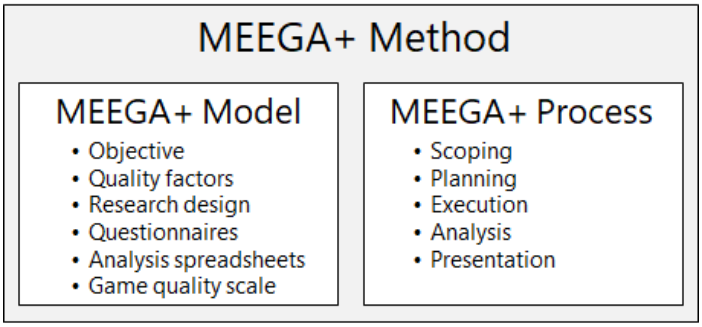

Figure 2. MEEGA+ Method

3.1 The MEEGA+ Model

The objective of the MEEGA+ model is to: evaluate the quality of educational games in terms of usability and player experience from the students' perspective in the context of computing education (Petri et al., 2018a).
Following the GQM approach (Basili et al., 1994), this objective is systematically decomposed into factors/dimensions to be measured. Based on the results of a mapping study (Petri et al., 2018b), analyzing conceptual definitions and similarities of the factors/dimensions, and on the results of a statistical analysis of the MEEGA+ model (exploratory factor analysis), analyzing a sample of 718 students from 40 case studies (Petri et al., 2018a; Petri et al., 2018b), we defined that the MEEGA+ model is decomposed into two quality factors and their dimensions (Figure 3). In this study, we define the usability as degree to which a product (educational game) can be used by specified users (students) to achieve specified goals with effectiveness and efficiency in a specified context of use (computing education), being composed of the following dimensions: aesthetics, learnability, operability, and accessibility

(ISO/IEC, 2014; Davis, 1989; Mohamed & Jaafar, 2010). And, the player experience is a quality factor that covers a deep involvement of the student in the gaming task, including its perception of learning, feelings, pleasures, and interactions with the game, environment and other players (Savi et al., 2011; O'Brien & Toms, 2010; Wiebe et al., 2014; Sweetser & Wyeth, 2005; Fu et al., 2009; Tullis & Albert, 2008; Keller, 1987; ISO/IEC, 2014; Savi et al., 2011; Sindre & Moody, 2003).

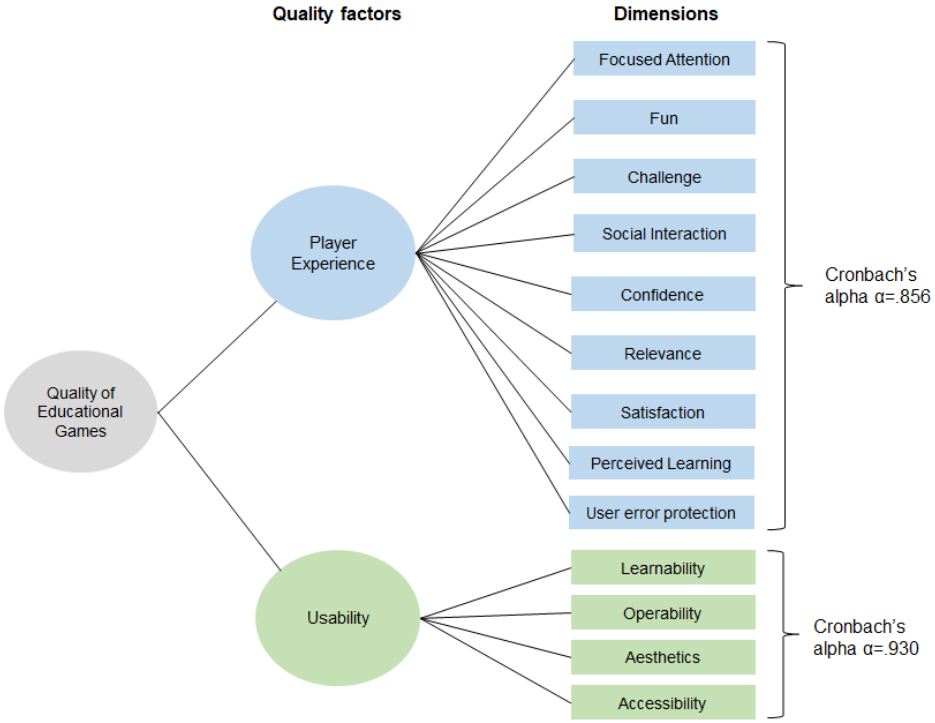

Figure 3. Decomposition of the MEEGA+ model

The definitions of these dimensions are presented in Table 1.

Table 1. Definition of the dimensions

Quality factor	Dimension	Definition
Usability	Aesthetics	Evaluating, if the game interface enables pleasing and satisfying interaction for the user (ISO/IEC, 2014).
	Learnability	Evaluating, if the game can be used by specified users to achieve specified goals of learning to use the game with effectiveness, efficiency, freedom from risk and satisfaction in a specified context of use (ISO/IEC, 2014).
	Operability	Evaluating the degree to which a game has attributes that make it easy to operate and control (ISO/IEC, 2014).
	Accessibility	Evaluating, if the game can be used by people with low/moderate visual impairment and/or color blindness (ISO/IEC, 2014).
Player experience	User error protection	Evaluating, if the game protects users against making errors (ISO/IEC, 2014; Fu et al., 2009). Applied only for evaluation of digital games.
	Focused Attention	Evaluating the attention, focused concentration, absorption and the temporal dissociation of the students (Keller, 1987; Wiebe et al., 2014; Savi et al., 2011).
	Fun	Evaluating the students' feeling of pleasure, happiness, relaxing and distraction (Poels et al., 2007, Savi et al., 2011).
	Challenge	Evaluating how much the game is sufficiently challenging with respect to the learner's competency level. The increase of difficulty should occur at an ap-

		propriate pace accompanying the learning curve. New obstacles and situations should be presented throughout the game to minimize fatigue and to keep the students interested (Sweetser & Wyeth, 2005; Savi et al., 2011).
	Social Interaction	Evaluating, if the game promotes a feeling of a shared environment and being connected with others in activities of cooperation or competition (Fu et al., 2009; Savi et al., 2011).
	Confidence	Evaluating, if students are able to make progress in the study of educational content through their effort and ability (e.g., through tasks with increasing level of difficulty) (Keller, 1987; Savi et al., 2011).
	Relevance	Evaluating, if students realize that the educational proposal is consistent with their goals and that they can link content with their professional or academic future (Keller, 1987; Savi et al., 2011).
	Satisfaction	Evaluating, if students feel that the dedicated effort results in learning (Keller, 1987; Savi et al., 2011).
	Perceived Learning	Evaluating the perceptions of the overall effect of the game on students' learning in the course (Sindre & Moody, 2003; Savi et al., 2011).

In order to operationalize the measurement of these defined quality factors/dimensions, a research design is defined.

Definition of the research design. As research design, we chose a case study design, which allows an in-depth research of an individual, group or event (Wohlin et al., 2012; Yin, 2017). The study is conducted as a one-shot post-test only design, in which the case study begins with the application of the treatment (educational game) and then the data are collected. Data collection is operationalized through a standardized measurement instrument (questionnaire). The questionnaire is answered by the students (self-assessment) in order to collect data on their perceptions about the game.

Definition of the MEEGA+ measurement instrument. In accordance with the research design and based on the defined factors/dimensions, we generate a set of items, improving the initial version of the MEEGA questionnaire, customizing and unifying existing standardized questionnaires (Savi et al., 2011; Keller, 1987; Tullis & Albert, 2008; Sindre & Moody, 2003; Sweetser & Wyeth, 2005; Poels et al., 2007; Gámez, 2009; Takatalo et al., 2010, O'Brien & Toms, 2010; Wiebe et al., 2014; Fu et al., 2009; Mohamed & Jaafar, 2010; Zaibon & Shiratuddin, 2010; Zaibon, 2015; Brooke, 1996; Davis, 1989). Table 2 shows the items of the MEEGA+ measurement instrument for each dimension/sub-dimension and their sources/references.

Table 2. MEEGA+ measurement instrument items and their references

Quality factor	Dimension	Item No.	Description
Usability	Aesthetics (ISO/IEC, 2014)	1	The game design is attractive (interface, graphics, cards, boards, etc.).
		2	The text font and colors are well blended and consistent.
	Learnability (ISO/IEC, 2014)	3	I needed to learn a few things before I could play the game.
		4	Learning to play this game was easy for me.
		5	I think that most people would learn to play this game very quickly.
	Operability (ISO/IEC, 2014)	6	I think that the game is easy to play.
		7	The game rules are clear and easy to understand.
	Accessibility (ISO/IEC, 2014)	8	The fonts (size and style) used in the game are easy to read.
		9	The colors used in the game are meaningful.
		10	The game allows customizing the appearance (font and/or color) according to my preferences.
Player experience	User error protection (Fu et al., 2009; ISO/IEC, 2014)	11	The game prevents me from making mistakes.
		12	When I make a mistake, it is easy to recover from it quickly.
	Confidence	13	When I first looked at the game, I had the impression that it would be easy for me.

	(Keller, 1987; Savi et al., 2011)	14	The contents and structure helped me to become confident that I would learn with this game.
	Challenge (Sweetser & Wyeth, 2005; Savi et al., 2011)	15	This game is appropriately challenging for me.
		16	The game provides new challenges (offers new obstacles, situations or variations) at an appropriate pace.
		17	The game does not become monotonous as it progresses (repetitive or boring tasks).
	Satisfaction (Keller, 1987; Savi et al., 2011)	18	Completing the game tasks gave me a satisfying feeling of accomplishment.
		19	It is due to my personal effort that I managed to advance in the game.
		20	I feel satisfied with the things that I learned from the game.
		21	I would recommend this game to my colleagues.
	Social Interaction (Fu et al., 2009; Savi et al., 2011)	22	I was able to interact with other players during the game.
		23	The game promotes cooperation and/or competition among the players.
		24	I felt good interacting with other players during the game.
	Fun (Poels et al., 2007, Savi et al., 2011)	25	I had fun with the game.
		26	Something happened during the game (game elements, competition, etc.) which made me smile.
	Focused Attention (Keller, 1987; Wiebe et al., 2014; Savi et al., 2011)	27	There was something interesting at the beginning of the game that captured my attention.
		28	I was so involved in my gaming task that I lost track of time.
		29	I forgot about my immediate surroundings while playing this game.
	Relevance (Keller, 1987; Savi et al., 2011)	30	The game contents are relevant to my interests.
		31	It is clear to me how the contents of the game are related to the course.
		32	This game is an adequate teaching method for this course.
		33	I prefer learning with this game to learning through other ways (e.g. other teaching methods).
	Perceived Learning (Sindre & Moody, 2003; Savi et al., 2011)	34	The game contributed to my learning in this course.
		35	The game allowed for efficient learning compared with other activities in the course.

The items presented in Table 2 compose the MEEGA+ measurement instrument in order to evaluate both digital and non-digital games for computing education. However, items 10, 11, and 12 are specifically used to evaluate digital games, related to the games' customization and user error protection. Thus, when evaluating non-digital games, these items are disregarded.

Items related to the learning goals of each game are included in the measurement instrument to be customized in accordance with the specific learning goals of each educational game. Typically, games for computing education are used to improve the knowledge on the cognitive levels of remembering, understanding, and application (ACM/IEEE-CS, 2013) in accordance with the revised version of Bloom's taxonomy (Anderson, Krathwohl, & Bloom, 2001). However, the MEEGA+ model is also flexible to cover goals related to higher cognitive levels, such as analyzing, evaluating and creating. Besides technical knowledge, games can also contribute to skill development, such as problem-solving, communication, teamwork, leadership, etc. (ACM/IEEE-CS, 2013). Such learning goals can be classified in accordance with the taxonomy of the psychomotor domain (Simpson, 1972). In addition, games can also contribute to develop professional attitudes, such as a mature behavior of the student, considering professional and legal issues as well as an ethical attitude in the profession (ACM/IEEE-CS, 2013). Learning goals related to this typically target the awareness and growth in attitudes, emotion, and feelings and can be classified in accordance to the taxonomy of affective domain (Krathwohl, Bloom, & Masia, 1973). Thus, for each learning goal of the game, the following statement should be customized in the MEEGA+ measurement instrument: The game contributed to <*verb related to the level of the learning goal (cognitive, psychomotor, and affective)*> <*goal/concept*>. For example, in accordance with the learning goals of SCRUMIA (Gresse von Wangenheim, Savi, & Borgatto, 2013), a game to reinforce the understanding of SCRUM concepts and to practice the SCRUM process, such statement

would be: "*The game contributed to recall concepts related to Sprint Planning*." These statements related to the learning goals of the game compose the MEEGA+ measurement instrument that is applied after the game session in order to capture the students' perceptions about their level of agreement (or disagreement) in the achieving of the learning objective(s) of the game.

Additionally, the MEEGA+ measurement instrument also supports the identification of the sample characteristics, collecting demographic information of the students that compose the sample of the study. Know the characteristics of the sample are important in order to interpret and analyse the data collected (Wohlin et al., 2012). Thus, the sample is characterized through the MEEGA+ measurement instrument in terms of age group, gender, and frequency that the students' play games (digital and/or non-digital games).

Response format. As response format, we adopt a 5-point Likert scale with response alternatives ranging from strongly disagree to strongly agree (DeVellis, 2016; Malhotra & Birks, 2008). The use of a Likert scale, in its original 5-point format, allows to express the opinion of the individual (student) under the object of study (educational game) with precision, besides allowing the individual being comfortable to express their opinion, using a neutral point and, thus, contributing to the quality of the answers (Dawes, 2008).

Analysis of the data collected. Following the MEEGA+ model, the objective defined is: evaluate the quality of educational games in terms of usability and player experience from the students' perspective in the context of computing education (Petri et al., 2018a).
Based on the objective defined for the evaluation, following the MEEGA+ model and the GQM approach (Basili et al., 1994), the objective is decomposed into quality aspects and analysis questions to be analysed:

Usability
AQ1: Does the <*name of the evaluated game*> has a good usability?
Player Experience
AQ2: Does the <*name of the evaluated game*> provides a positive player experience?

In addition to this analysis questions, complementary questions may address the identification of the characteristics of the sample in terms of age, gender and frequency that the students play games:
AQ3: How old are the students that compose the sample of the study?
AQ4: What is the gender of the students that compose the sample of the study?
AQ5: What is the frequency that the students play digital and/or non-digital games?

In order to answer these defined analysis questions, data collected through the MEEGA+ measurement instrument are analysed through descriptive statistics methods (Trochim & Donnelly, 2008; Wohlin et al., 2012). Descriptive statistics methods are used to describe and graphically present interesting aspects of the data collected (Wohlin et al., 2012). Table 3 presents the descriptive statistics methods used to answer each analysis question.

Table 3. Descriptive statistics methods used to answer the analysis questions

Analysis questions	Descriptive statistics methods	Data analysed
AQ1: Does the <*name of the evaluated game*> has a good usability?	Measures of central tendency (median, average and frequency of responses); Graphical visualization (frequency charts)	Data collected through the MEEGA+ measurement instrument on the usability quality factor.
AQ2: Does the <*name of the evaluated game*> provides a posi-	Measures of central tendency (median, average and frequency	Data collected through the MEEGA+ measurement instrument on the player

tive player experience?	of responses); Graphical visualization (frequency charts)	experience quality factor.
AQ3: How old are the students that compose the sample of the study?	Measures of central tendency (frequency of responses); Graphical visualization (frequency charts)	Data collected through the MEEGA+ measurement instrument on demographic information of the sample.
AQ4: What is the gender of the students that compose the sample of the study?	Measures of central tendency (frequency of responses); Graphical visualization (frequency charts)	Data collected through the MEEGA+ measurement instrument on demographic information of the sample.
AQ5: What is the frequency that the students play digital and/or non-digital games?	Measures of central tendency (frequency of responses); Graphical visualization (frequency charts)	Data collected through the MEEGA+ measurement instrument on demographic information of the sample.

In order to operationalize the analysis of the data collected and assist researchers to answer the defined analysis questions using descriptive statistics methods, a standardized analysis spreadsheet was prepared using Microsoft Excel (desktop version). The spreadsheet contains five internal spreadsheets (Data; Demographic data; Graphs, Demographic graphs; and Items).

In order to answer the analysis questions AQ1 and AQ2, frequency charts are automatically generated in the analysis spreadsheet, presenting the frequency of the responses and the median for each measurement instrument item. The graphics are presented in accordance with the quality factors and dimensions of the MEEGA+ model (Figure 4).

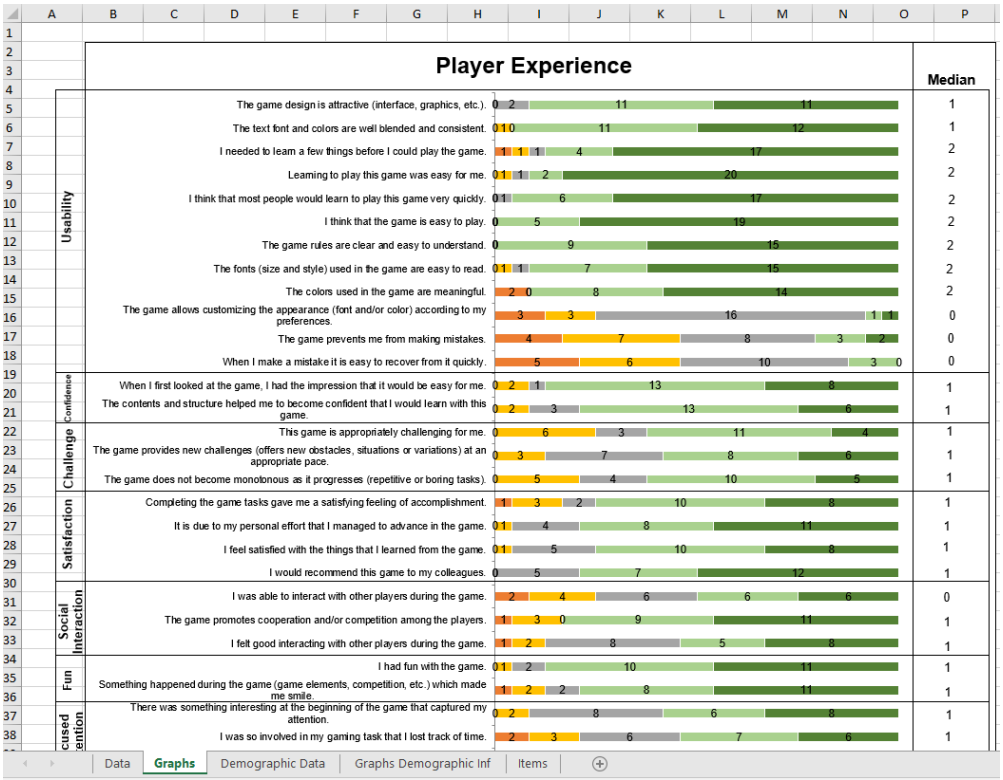

Figure 4. Graphs spreadsheet

Similarly, the graphics presenting the demographic information are generated and presented in the Demographic graphs spreadsheet (Figure 5), providing information to answer the analysis questions AQ3, AQ4, and AQ5.

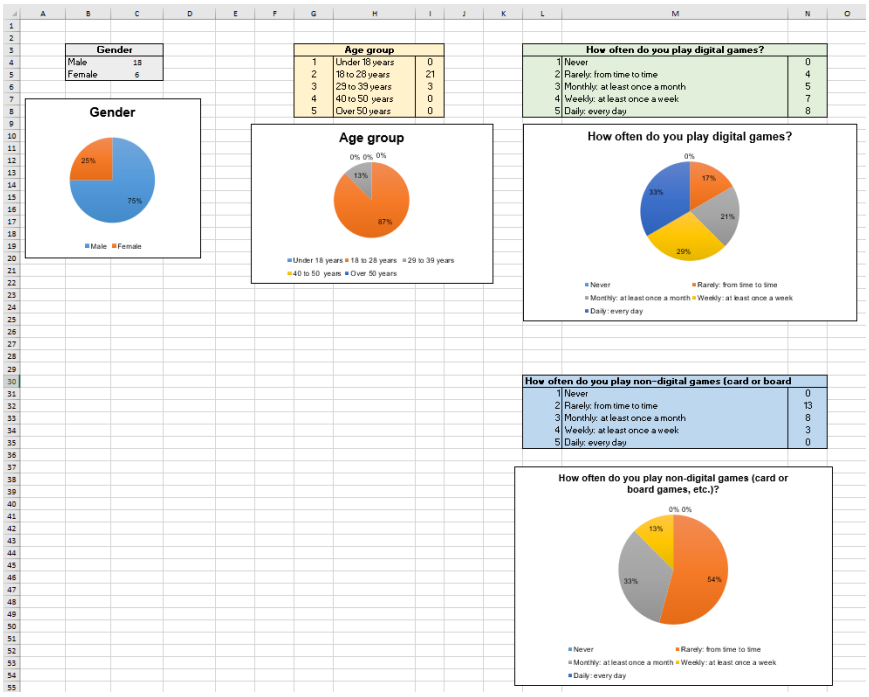

Figure 5. Demographic graphs spreadsheet

In order to assist instructors/researchers in the use of the MEEGA+ data analysis spreadsheet, an instruction guide was prepared. This guide describes the step by step for the use of the MEEGA+ data analysis spreadsheet, describing how the data should be inserted and prepared to support the automatic generation of the descriptive analysis.

Game Quality Scale

In order to classify the evaluated game through the MEEGA+ model on a quality level, a MEEGA+ scale has been defined. The MEEGA+ scale classifies the evaluated game regarding to its quality level (low, good, or excellent quality), based on students' perception. The MEEGA+ scale was developed by adopting the statistical technique of the Item Response Theory (IRT) (Pasquali & Primi, 2003), applying mathematical models, and defining three quality levels, based on answers of 1031 students (Table 4).
The quality level of the evaluated game is determined based on the data collected using the MEEGA+ measurement instrument and analysing them through an R script which applies the defined scale scores in the collected data.

Table 4. Game quality levels

Quality level	Level description
Low quality (θ < 42.5)	At this level, the game rarely provides social interaction and hardly ever produces moments of fun among the players. The game does not capture the student's focused attention, does not arouse the confidence that he/she will learn from the game, nor does it produce feelings of satisfaction. The game rarely presents challenges, has monotonous tasks and does not contribute to student learning. Although a game at this level has a low relevance to the students' interests, a student recognizes that the game's content is related to the course. In terms of usability, a game at this level sometimes exhibits operability features, which may have some clear rules and be easy to play.
Good quality	At this level, the game sometimes presents challenging activities, offering

(42.5 <= θ < 65)	new challenges for students. It provides moderately focused attention to the players, although students do not forget about their surroundings. Sometimes the game also provides feelings of confidence and satisfaction in the players. Frequently the game presents moments of social interaction and fun among the players. Often the game is considered relevant to the students' interests and, usually, the students recognize that the game's content is related to the course. Frequently the game contributes efficiently to student learning. In terms of usability, the game usually has the clear rules and is easy to play, although, usually does not present a fully attractive design.
Excellent quality (θ >= 65)	At this level, the game is challenging for students and has no monotonous activities. It is highly relevant to students' interests and provides excellent focused attention, satisfaction, fun, and social interaction. It allows the student to be confident that he/she will learn from the game and contribute to and efficient student learning. In terms of usability, the game presents excellent operability and learnability, that is, it has clear rules and is easy to learn to play. Even so, a game at this level may present improvements in terms of aesthetics, not presenting a fully attractive design.

A complete description of the MEEGA+ scale can be found in Petri et al. (2018c).

3.2 The MEEGA+ Process

In order to guide the application of the model, the MEEGA+ method also contains a systematic process for the execution of the evaluation of educational games. The process provides a systematic support, detailing steps and interrelated activities, guiding researchers and instructors in the conduction of game evaluations.
The MEEGA+ process has been designed in a prescriptive way, defining how the process should be executed, establishing rules and procedures (Acuña & Ferré, 2001). Thus, the MEEGA+ process is described through phases, activities and work products (Acuña et al., 2000; Benali & Derniame, 1992; Finkelstein et al., 1994):

- **Phase** is a set of activities grouped in steps, presenting a logical and structured sequence.
- **Activity** is the stage of a process that produces externally visible changes of state in the product. An activity can have inputs, outputs, intermediate results (work products). The activity implements procedures, rules and objectives to transform a product.
- **Work products** are the inputs and outputs of an activity from a process. They can be produced and consumed throughout the process and can have long life cycles, being created, accessed and modified (changing its status) during the process. Each work product has a status. The status of a work product indicates its position in the process, determining which actions can be performed on the work product (Münch et al., 2012). Table 5 presents the status of the work products adopted in the MEEGA+ process, defined based in Münch et al. (2012).

Table 5. Status of the work products

Status	Description
Initial	Work products are initially set to this status. Typically, there is not an action taken on the work product. Typically, a work product with this status is a work product created in the process or an existing product that needs a customization, used as an input to an activity such as, a game developed, a template of a document/spreadsheet, etc. Example: MEEGA+ data collection instruments
In Preparation	Indicating that the work product is incomplete, not ready for use. Examples: the game needs to be installed, the data collections instruments need to be printed.

Ready to use	Indicating that all activities associated with the work product has been completed and the work product is ready to be used. Examples: the game is installed and ready to play, the questionnaires are printed.

Phases and activities of the MEEGA+ process are defined based on the process of empirical studies proposed by Wohlin et al. (2012), and in practical experiences of our research group (GQS/INCoD/INE/UFSC) in conduction of game evaluations (http://www.gqs.ufsc.br/software-engineering-education/). Figure 6 presents the MEEGA+ process modeling adopting the BPM notation.

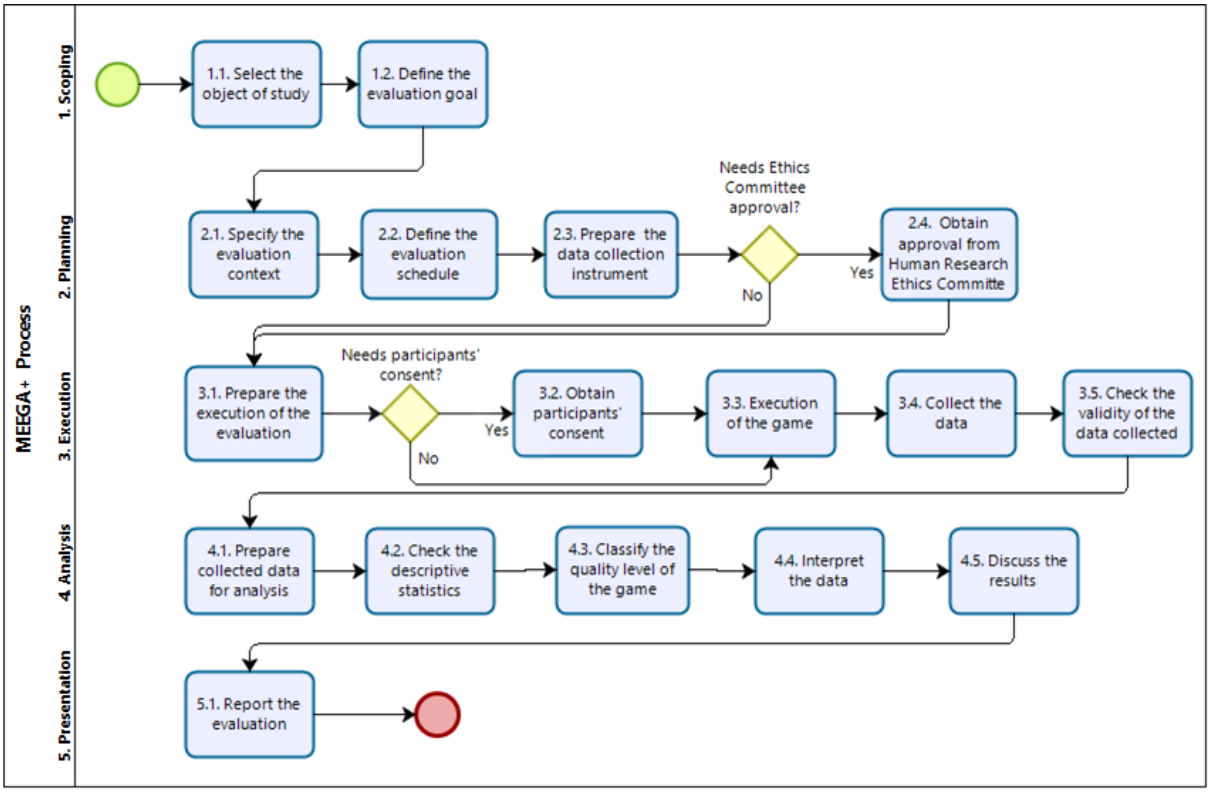

Figure 6. **The MEEGA+ Process**

The MEEGA+ Process may be conducted either manually supported by tools, using the standardized questionnaires, spreadsheets and scripts provided by the MEEGA+ Model (available in English, Spanish, and Brazilian Portuguese at http://www.gqs.ufsc.br/meega-a-model-for-evaluating-educational-games/). Or semi-automated, adopting the AssistantMEEGA+ v0.9 (http://assistantmeega.com.br/), developed by our research partners from USES/UFAM (Grupo de Usabilidade e Engenharia de Software, Universidade Federal de Amazonas). However, the current version (v0.9) of the AssistantMEEGA+ provides only a partial support to the MEEGA+ Process, not supporting all activities of the MEEGA+ Process, e.g. not supporting the classification of the quality level of the evaluated game, using the MEEGA+ scale.

Typically, game evaluations are conducted by researchers. However, the MEEGA+ process is described in order to be applied by any student or instructor, not requiring advanced knowledge in education, measurement or statistics. Thus, in this process, we use the generic term researcher representing any student/instructor/etc. responsible for the game evaluation.

The MEEGA+ Process is organized into phases and activities:

Phase 1. Scoping. In the first phase of the MEEGA+ process, the evaluation scope is defined. The scope of the evaluation is set by defining its goals and the object of study (Table 6).

Table 6. Activities of the Scoping phase

Phase 1. Scoping	
Activity 1.1. Select the object of study (educational game)	
Description	The objective of this activity is to select the object of study (educational game). An educational game is classified as an instructional strategy that typically involves competition and is organized by rules and restrictions to achieve a certain educational goal. Typically, an educational game is characterized by various elements, such as goals, rules, restrictions, interaction, challenge, competition, rewards and feedback (Abt, 2002; Prensky, 2007; Ritterfeld et al., 2010; Djaouti et al., 2011). The selected game may be a digital game (computer game) or a non-digital game (board, cards game, etc.). Digital games are electronic games that involve human interaction with a user interface to generate visual feedback on an electronic device such as smartphones, computers, tablets, etc., whereas non-digital games are played with non-digital resources such as game boards, cards, pencils and papers, etc. (Prensky, 2007; Ritterfeld et al., 2010; Djaouti et al., 2011). The selected game should be specific to the development of a skill, knowledge, or attitude, for any knowledge area in the context of computing education. As a suggestion, the game may be selected from game repositories: https://pmiteach.org/teaching-pm/resources/ https://www.infoq.com/news/2008/10/agile-games https://list.ly/list/CL-management-games-and-simulations-for-itsm http://games-factory-online.nl/seriousgames-english/seriousgamescatalogue/ http://www.semq.eu/leng/proimplo.htm http://tastycupcakes.org/pt http://www.gqs.ufsc.br/igr In addition, the game may also be selected from the results of systematic reviews that report several games for computing education: Battistella, P.; Gresse von Wangenheim, C. Games for Teaching Computing in Higher Education – A Systematic Review. IEEE Technology and Engineering Education (ITEE) Journal, 9(1), 8-30, 2016. Available at: http://www.gqs.ufsc.br/wp-content/uploads/2011/11/ITEE-Games-for-Teaching-Computing-in-Higher-Education_Vdraft.pdf Petri, G.; Gresse von Wangenheim, C. How games for computing education are evaluated? A systematic literature review. Computers & Education, 107, 68-90, 2017. Available at: http://dx.doi.org/10.1016/j.compedu.2017.01.004
Premise	As this process does not support the development of a new game, the selected game must have already been developed. It is expected that the selected game will not be offensive or violent. Does not discriminate against students in terms of race, sex, religion, nationality, disability, sexual orientation or age. And, it should attend the game's conditions of use (copyright).
Work products	**Input:** P1.1.1 – Educational game [Initial]
	Output: P1.1.1 – Educational game [In Preparation]
Activity 1.2. Define the evaluation goal	

Description	This activity aims to define the evaluation goal in terms of the object of study, purpose, quality aspects, perspective, and context. The evaluation goal is defined considering the definitions proposed by Wohlin et al. (2012) and Basili et al. (1994): ▪ The object of study is the entity that is analysed in the evaluation. ▪ The purpose defines what the intention of the evaluation. ▪ The quality aspects define the factors under analysis in the evaluation. ▪ The perspective tells the viewpoint from which the evaluation results are interpreted. ▪ The context is the environment in which the evaluation is run. Thus, based on the definition of the MEEGA+ model, which follows the GQM goal template (Basili et al., 1994), the evaluation objective is defined: Analyse the <*name of the selected* game> for the purpose of *evaluate the quality* in terms of *usability and player experience* from the *students' point of view* in the context of *higher computing education*.
Work products	**Input:** P1.1.1 – Educational Game [In Preparation]; P1.2.1 – MEEGA+ evaluation goal [Ready to use] The detailed description of the MEEGA+ model is available at <http://www.gqs.ufsc.br/meega-a-model-for-evaluating-educational-games/>
	Output: P1.2.2 - Evaluation goal [Ready to use]

Phase 2. Planning. After scoping of the evaluation, it is planned (Table 7).

Table 7. Activities of the Planning phase

Phase 2. Planning	
Activity 2.1. Specify the evaluation context	
Description	This activity aims to specify the context where the evaluation will be conducted. Typically, an educational game is used in an educational context. Thus, in this activity, the context is specified in terms of educational institution, program and course, in which the game will be applied.
Work products	**Input:** P1.1.1 – Educational Game [In Preparation]; P1.2.2 - Evaluation goal [Ready to use]
	Output: P2.1.1 – Evaluation context [Ready to use]
Activity 2.2. Define the evaluation schedule	
Description	This activity aims to plan the evaluation schedule, defining the date, hour and place of the application of the game in the defined course. If the selected game is a digital game and needs a computer lab (with computers, tablets, etc.) available to play it. In this activity, the lab must be defined and scheduled, considering the expected population (P2.1.1 – Evaluation context).
Work products	**Input:** P1.1.1 – Educational Game [In Preparation]; P2.1.1 – Evaluation context [Ready to use]
	Output: P2.2.1 – Evaluation schedule [Ready to use]
Activity 2.3. Prepare the data collection instrument	
Description	This activity aims to prepare/customize the data collection instrument, provided by the MEEGA+ Model, for the evaluation of the selected game. If the evaluation is conducted manually, it is necessary to choose the questionnaire in accordance with the game's platform (digital or non-digital), language (English, Span-

ish, and Brazilian Portuguese).

Information about the evaluation context (P2.1.1 – Evaluation context) can be previously inserted in the questionnaire in the Demographic Information section, including the game's name, researcher, place and date, institution, program, and course.

Figure 7 presents an example based in the non-digital game SCRUMIA (Gresse von Wangenheim, Savi, & Borgatto, 2013):

Questionnaire for quality evaluation of non-digital games

Game's name: SCRUMIA

Please, help us improve the game answering the following questions. All information is collected anonymously and will be used only in a summarized way in the context of this game evaluation.

Instructor name: _____.
Place and date: _____.

Demographic Information	
Institution:	
Undergraduate program:	
Course:	
Age group:	☐ Under 18 years ☐ 18 to 28 years ☐ 29 to 39 years ☐ 40 to 50 years ☐ Over 50 years
Gender:	☐ Male ☐ Female
How often do you play digital games?	☐ Never ☐ Rarely: from time to time ☐ Monthly: at least once a month ☐ Weekly: at least once a week ☐ Daily: every day.
How often do you play non-digital games (card or board games, etc.)?	☐ Never ☐ Rarely: from time to time ☐ Monthly: at least once a month ☐ Weekly: at least once a week ☐ Daily: every day.

Figure 7. **The MEEGA+ questionnaire**

Another customization is in terms of the learning objectives of the selected game. For each learning objective of the game, the following statement should be customized in the questionnaire:

The game contributed to <*verb related to the level of the learning goal (cognitive, psychomotor, and affective)*> <goal/concept>. An example, in accordance with the learning goals of SCRUMIA (Gresse von Wangenheim, Savi, & Borgatto, 2013), would be:

"The game contributed to recall concepts related to Sprint Planning."

Figure 8 presents an example of the statement regarding the learning objective of SCRUMIA.

Statements	Select an option as your evaluation				
	Strongly disagree	Disagree	Neither disagree nor agree	Agree	Strongly agree
The game contributed to my learning in this course.	☐	☐	☐	☐	☐
The game allowed for efficient learning compared with other activities in the course	☐	☐	☐	☐	☐
The game contributed to *recall* the concepts related to *Sprint Planning*.	☐	☐	☐	☐	☐

Figure 8. **The MEEGA+ questionnaire**

The MEEGA+ method is also flexible to include additional items if the researcher needs to evaluate another quality factor(s). However, as this factor is outside the scope of the MEEGA+, it is not possible to assure its validity and reliability in the measurement of new factors included.

Work	**Input:** P1.1.1 – Educational Game [In Preparation]; P2.3.1 – MEEGA+ data collection

products	instruments [Initial]; P2.1.1 – Evaluation context [Ready to use]; P2.2.1 – Evaluation schedule [Ready to use] The MEEGA+ questionnaires are available in English, Spanish, and Brazilian Portuguese at http://www.gqs.ufsc.br/meega-a-model-for-evaluating-educational-games/.
	Output: P2.3.1 – MEEGA+ data collection instruments [In Preparation]
Activity 2.4. Obtain approval from the Human Research Ethics Committee (optional)	
Description	Although the evaluation of a game offers minimal risk to the participants, some educational institutions require that all research involving humans be approved by the ethics committee. Thus, before the conduction of the evaluation is necessary verify the requirements of the ethics committee of the institution that the evaluation will be conducted. To approve a research involving humans, an ethics committee, typically, requires the declaration of a coordinator, a research project, data collection instruments, and a consent form.
Work products	**Input:** P1.1.1 – Educational Game [In Preparation]; P2.3.1 – MEEGA+ data collection instruments [In Preparation]; P2.1.1 – Evaluation context [Ready to use]; P2.2.1 – Evaluation schedule [Ready to use]
	Output: P2.4.1 – Approval of the Ethics Committee (optional) [Ready to use]

Phase 3. Execution. When an evaluation has been planned, it is carried out in order to collect the data to be analysed. This third phase aims to organize and define the execution of the game evaluation for the selected participants. In this phase, data are collected in order to achieve the evaluation objective (Table 8).

Table 8. Activities of the Execution phase

Phase 3. Execution	
Activity 3.1. Prepare the execution of the evaluation	
Description	This activity aims to prepare the materials required for the execution of the game. In this activity, if necessary, the data collection instruments and consent forms should be printed, based on the number of the participants described in P2.1.1 – Evaluation context. To non-digital games: organize and/or print (if necessary) the materials of non-digital games, such as boards, cards, etc. To digital games: organize and install (if necessary) the game. Additional materials should be prepared, such as projector, mobile devices for digital games. And, tables and chairs for non-digital games.
Work products	**Input:** P1.1.1 – Educational Game [In Preparation]; P2.1.1 – Evaluation context [Ready to use]; P2.2.1 – Evaluation schedule [Ready to use]; P2.3.1 – MEEGA+ data collection instruments [In Preparation]
	Output: P2.3.1 – MEEGA+ data collection instruments [Ready to use]
	Output: P3.1.2 – Consent form (optional) [Ready to use]
	Output: P1.1.1 – Educational Game [Ready to use]
Activity 3.2. Obtain participants' consent (optional)	
Description	Before the execution of the game, if necessary, the consent form should be signed by all participants, indicating that they agree and accept to participate in the research.
Work products	**Input:** P3.1.2 – Consent form (optional) [Ready to use]
	Output: P3.2.1 – Participants' consent (optional) [Ready to use]
Activity 3.3. Execution of the game	
Description	During this activity the game is applied in the participants, using the game materials (P3.1.3 – Game materials) and considering the evaluation schedule (P2.2.1 – Evaluation schedule) and context (P2.1.1 – Evaluation context). Before the game execution, the instructor presents the game and explain its rules. Then, the participants can start playing the game.

Work products	**Input:** P1.1.1 – Educational Game [Ready to use]; P2.1.1 – Evaluation context [Ready to use]; P2.2.1 – Evaluation schedule [Ready to use]; P2.3.1 – MEEGA+ data collection instruments [Ready to use]
	Output: P3.3.1 – Game executed [Ready to use]
Activity 3.4. Collect the data	
Description	After the execution of the game, data collection takes place. If the evaluation is being conducted using the AssistantMEEGA+, this activity is performed through an electronic form of the AssistantMEEGA+. If the evaluation is being conducted manually, the data collection instrument should be distributed to the participants in order to them fill out the questionnaire based on their perceptions about the game.
Work products	**Input:** P2.3.1 – MEEGA+ data collection instruments [Ready to use]; P3.3.1 – Game executed [Ready to use]
	Output: P3.4.1 – Data collected [Initial]
Activity 3.5. Check the validity of the data collected	
Description	When data has been collected, the researcher must check if the data is reasonable and that it has been collected correctly. This covers aspects such as if the participants have understood the items and therefore answered them correctly. Thus, based on Wohlin et al. (2012), Yin (2017) and Izquierdo & Pedrero (2014), criteria are defined in order to check if the data collected are valid: ▪ Check if all participants have participated seriously of the evaluation (e.g. not answering all the questionnaire items using the same response category). ▪ Check if all questionnaire items have only one response (one response category). ▪ Check if each questionnaire was answered completely and correctly (not missing to answer more than 10% of the questionnaire items (4 or more items) (Izquierdo & Pedrero, 2014)). Questionnaires answered incompletely (missing to answer 4 or more questionnaire item) and that not meet the defined criteria, should be disregarded of the analysis and the participant removed of the sample.
Work products	**Input:** P3.4.1 – Data collected [Initial]
	Output: P3.4.1 – Data collected [In Preparation]

Phase 4. Analysis. After collecting data in the execution phase, conclusions are drawn based on this data. To be able to draw valid conclusions, evaluation data must to be analysed. Thus, the fourth phase of the MEEGA+ process aims to interpret and analyse the data collected in the execution phase (Table 9).

Table 9. Activities of the Analysis phase

Phase 4. Analysis	
Activity 4.1. Prepare collected data for analysis	
Description	If the evaluation was conducted using the AssistantMEEGA+, this activity is performed in an automatic way. If the evaluation was conducted manually, the data collected using the MEEGA+ questionnaire must be prepared for the data analysis process by introducing them in the data analysis spreadsheet provided by the MEEGA+ model. Select and download the spreadsheet specifically for the kind of game platform (digital or non-digital). The researcher needs introduce the data collected in the spreadsheet and customize the learning objectives, following the detailed instructions for the use of the MEEGA+ data analysis spreadsheet. As presented in Figure 9, each row of the data analysis spreadsheet represents an answer of one student for the game evaluation (one questionnaire). The columns represent the MEEGA+ measurement instrument items and are organized in the spreadsheet in the same order that are presented in the questionnaire.

Each answer must be introduced in the spreadsheet, following the definition as presented in Table 9.1.

Table 9.1. Introducing an answer in the spreadsheet

Answer	Introduce into the spreadsheet
Strongly disagree	-2
Disagree	-1
Neither disagree nor agree	0
Agree	1
Strongly Agree	2

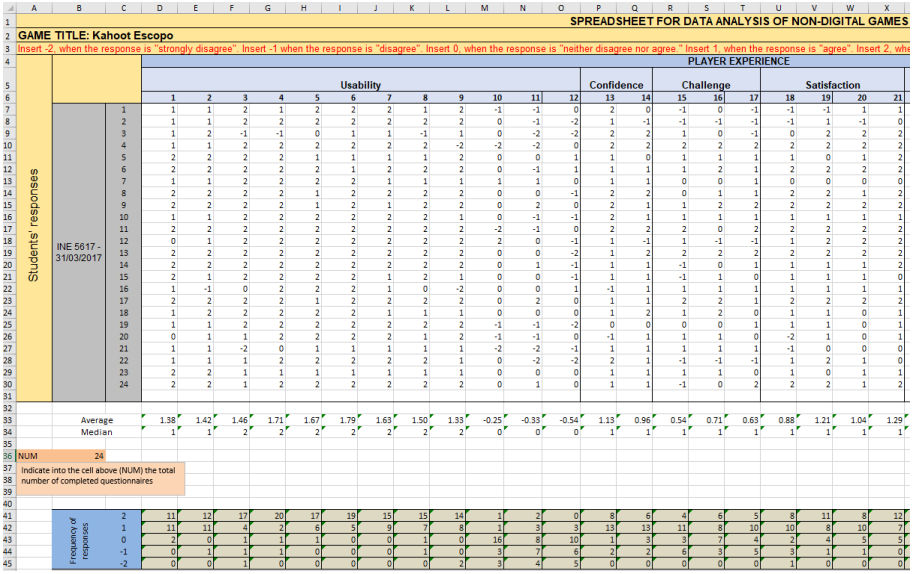

Figure 9. **Data analysis spreadsheet**

Work products	**Input:** P3.4.1 – Data collected [In Preparation]; P4.1.1 – MEEGA+ data analysis spreadsheets and instructions guide [Initial]; The spreadsheets are available in English, Spanish, and Brazilian Portuguese at <http://www.gqs.ufsc.br/meega-a-model-for-evaluating-educational-games/> Instructions also available in English, Spanish, and Brazilian Portuguese at <http://www.gqs.ufsc.br/meega-a-model-for-evaluating-educational-games/>
	Output: P3.4.1 – Data collected [Ready to use]
	Output: P4.1.1 – MEEGA+ data analysis spreadsheets [In Preparation]
Activity 4.2. Check the descriptive statistics	
Description	Descriptive statistics deal with the presentation and numerical processing of a data set. After the data collected are prepared and organized in the analysis spreadsheets, the data are characterized by automatically calculating descriptive statistics. Thus, in this activity it is necessary only to check if the descriptive statistics were generated correctly. The spreadsheet automatically calculates the frequency of responses, average and median for each measurement instrument item, as presented in Figure 10.

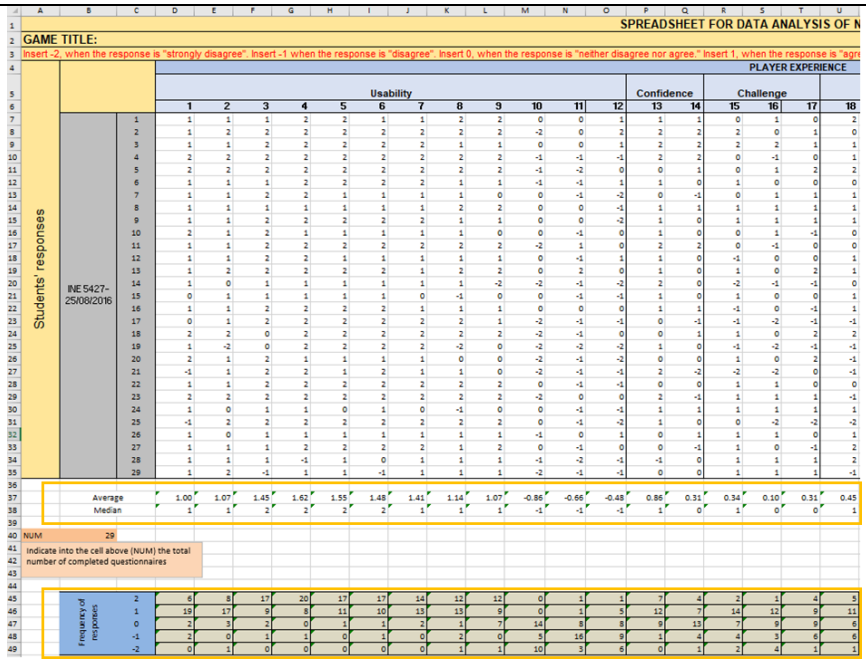

Figure 10. **Data analysis spreadsheet**

Graphics of frequency are also automatically generated and presented in the Graphs spreadsheet (Figure 11). The graphics are presented in accordance with the quality factors and dimensions of the MEEGA+ model.

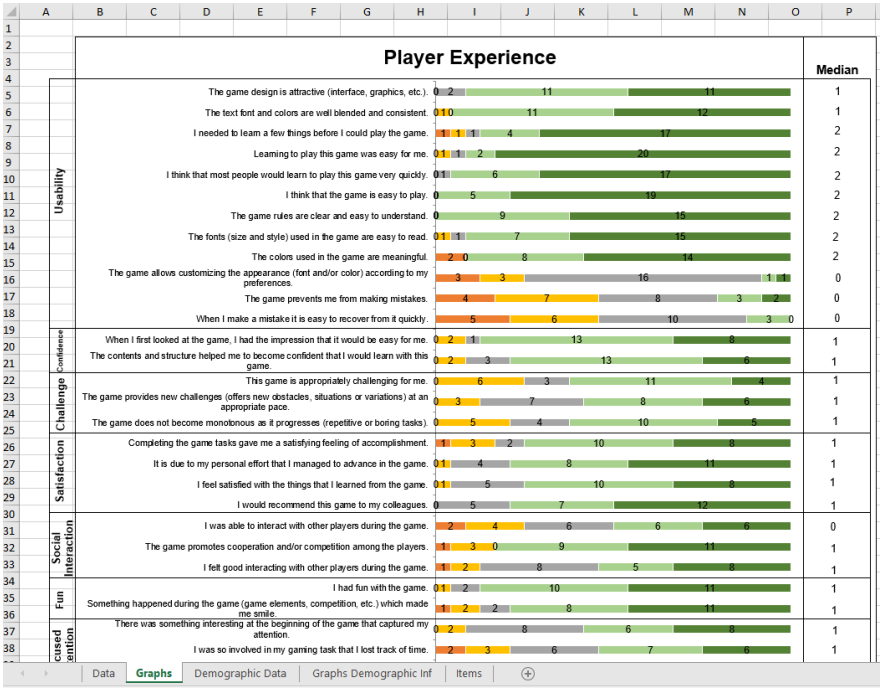

Figure 11. **Frequency graphs**

Similarly, the graphics presenting the demographic information are generated and presented in the Demographic graphs spreadsheet (Figure 12).

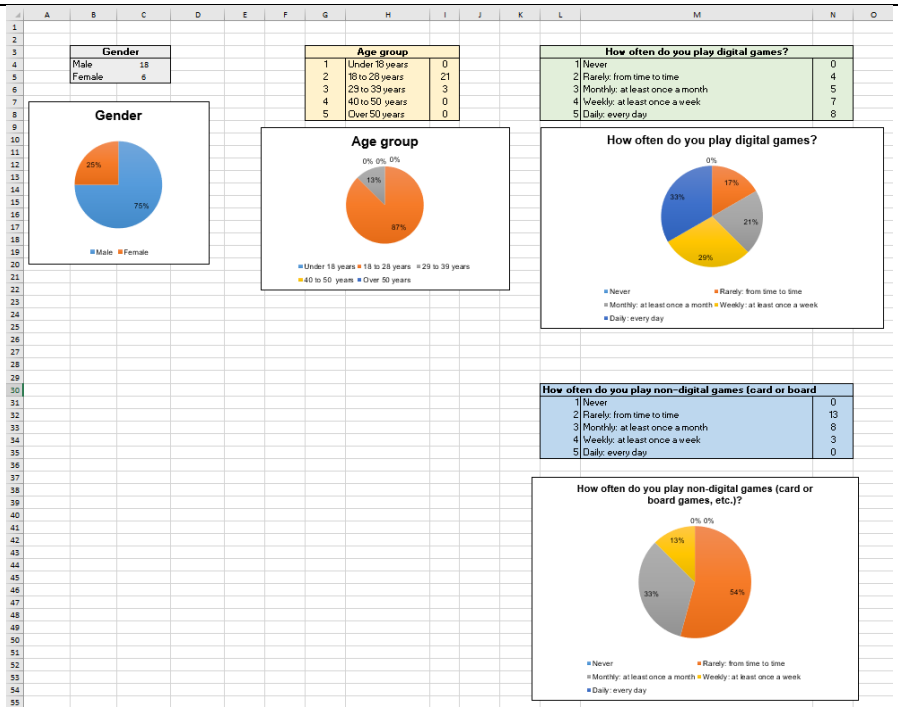

Figure 12. **Demographic information graphs**

Work products	**Input:** P3.4.1 – Data collected [Ready to use]; P4.1.1 – MEEGA+ data analysis spreadsheets [In Preparation]
	Output: P4.2.1 – Descriptive statistics results [Ready to use]
	Output: P4.1.1 – MEEGA+ data analysis spreadsheets [Ready to use]
Activity 4.3. Classify the quality level of the game	
Description	If the evaluation was conducted using the AssistantMEEGA+, this activity is not supported in the current version (v0.9). Thus, this activity needs to be performed manually. After the data collected are prepared and organized in the analysis spreadsheets, these data may be used to determine the quality level of the evaluated game, classifying it based on the game quality scale, defined in the MEEGA+ model. In order to classify the game using the MEEGA+ scale, it is necessary to follow these steps: **1. Download the files**: it is necessary to download the files of the scale (available at: http://www.gqs.ufsc.br/meega-a-model-for-evaluating-educational-games/) and organize them into a single directory on your computer. **2. Prepare the files**: Prepare an auxiliary file (named EXTRA) in csv (comma-separated values) extension with the data collected for the items (1, 5, 6, 7, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35). This file is an input for an R script which applies the defined scale scores in the data collected. This auxiliary file must be in the same directory that the other files. **3. Use RStudio:** Using the statistical software RStudio (https://www.rstudio.com/), load the defined directory and execute the R script (SCRIPT_TRI_SCORE), which calculates the scores of the data collected, applying the Item Response Theory. As a result, on the console of the RStudio, the scores will be presented, and a new auxiliary file (SCORE_TRI_EXTRA.csv) will be generated in the directory, presenting the calculated scores and the standard error, as shown the Figure 13.

	A	B	C	D
1		ID	SCORE_TRI	ERRO_PADRAO_TRI
2	1	N1	-1,3648	0,6962
3	2	N2	-1,351	0,6132
4	3	N3	-0,8182	0,5252
5	4	N4	1,836	0,3176
6	5	N5	1,1933	0,3155

Figure 13. **TRI scores**

4. Analyse the scores and classifying the game: The Item Response Theory calculates the score (column SCORE_TRI) (in a (0,1) scale) of an individual and positions it on the defined scale. However, we are interested in the classification of the game (and not of an individual). Thus, we must calculate the average of the provided scores (column SCORE_TRI) of all participants that evaluated the game.

Based on the scores presented in Figure 17, the average for these scores is θ=-0.10. However, in order to provide a better understanding of these values, we transform this scores in a (50,15) scale, applying the following formula $\theta_{50,15}=50+15*\theta_{0,1}$. Thus, applying this formula in the average score we obtained a value of θ=48.5.

Based on this final value, we may classify the evaluated game in the MEEGA+ game scale. With a score of θ=48.5, this game is classified as a game with good quality (42.5 <= θ < 65). Therefore, this game typically presents the characteristics of its quality level, as described in Table 9.2.

Table 9.2. Game quality levels

Quality level	Level description
Low quality (θ < 42.5)	At this level, the game rarely provides social interaction and hardly ever produces moments of fun among the players. The game does not capture the student's focused attention, does not arouse the confidence that he/she will learn from the game, nor does it produce feelings of satisfaction. The game rarely presents challenges, has monotonous tasks and does not contribute to student learning. Although a game at this level has a low relevance to the students' interests, a student recognizes that the game's content is related to the course. In terms of usability, a game at this level sometimes exhibits operability features, which may have some clear rules and be easy to play.
Good quality (42.5 <= θ < 65)	At this level, the game sometimes presents challenging activities, offering new challenges for students. It provides moderately focused attention to the players, although students do not forget about their surroundings. Sometimes the game also provides feelings of confidence and satisfaction in the players. Frequently the game presents moments of social interaction and fun among the players. Often the game is considered relevant to the students' interests and, usually, the students recognize that the game's content is related to the course. Frequently the game contributes efficiently to student learning. In terms of usability, the game usually has the clear rules and is easy to play, although, usually does not present a fully attractive design.
Excellent quality (θ >= 65)	At this level, the game is challenging for students and has no monotonous activities. It is highly relevant to students' interests and provides excellent focused attention, satisfaction, fun, and social interaction. It allows the student to be confident that he/she will learn from the game and contribute to and efficient student learning. In terms of usability, the game presents excellent operability and learnability, that is, it has clear rules and is easy to learn to play. Even so, a game at this level may present improvements in terms of aesthetics, not presenting a fully attractive design.

Work products	**Input:** P3.4.1 – Data collected [Ready to use]; P4.1.1 – MEEGA+ data analysis spreadsheets [Ready to use]; P4.3.1 – MEEGA+ game quality scale [Ready to use] The files of the scale are available at http://www.gqs.ufsc.br/meega-a-model-for-evaluating-educational-games/.
	Output: P4.3.2 – Game quality level [Ready to use]
Activity 4.4. Interpret the data	
Description	When the data collected are organized and characterized by descriptive statistics, it is important interpret them in order to achieve the evaluation goal, defined in the scoping phase. The analysis is performed following the quality aspects determined in the evaluation goal. Following the MEEGA+ model, the game is evaluated in terms of usability and

	player experience. Thus, we suggest that the data be analysed for each quality factor (usability and player experience). For each quality factor, analyses the frequency of each measurement instrument item, identifying the degree of agreement and/or disagreement of the students. Based on this degree of agreement and/or disagreement, strengths and weaknesses of the evaluated game may be also identified and reported. In addition, the quality level of the evaluated game, defined by the MEEGA+ scale, may be reported, describing the characteristics of the game in respect of its quality level.
Work products	**Input:** P3.4.1 – Data collected [Ready to use]; P4.1.1 – MEEGA+ data analysis spreadsheets [Ready to use]; P4.2.1 – Descriptive statistics results [Ready to use]; P4.3.2 – Game quality level [Ready to use]; P1.2.2 - Evaluation goal [Ready to use]
	Output: P4.4.1 – Evaluation results [Ready to use]
Activity 4.5. Discuss the results	
Description	This activity aims to discuss the findings identified in the evaluation results, indicating the main contribution of the use of this game as an instructional strategy for computing education, as well as its improvement opportunities. In addition, the results of the evaluated game may be compared to results of related studies and game evaluations, indicating the similarities and differences to other games that have similar learning objectives. Furthermore, it is important identifying threats to the study validity, as well as report mitigation strategies adopted in order to minimize the impact in the study.
Work products	**Input:** P4.4.1 – Evaluation results [Ready to use]
	Output: P4.5.1 – Discussion [Ready to use]

Phase 5. Presentation. In this phase the evaluation is reported (Table 10).

Table 10. Activities of the Presentation phase

Phase 5. Presentation	
Activity 5.1. Report the evaluation	
Description	This activity aims to produce an evaluation report describing, in detail, how the evaluation of the selected game was defined, planned, executed and analysed. We focus on the academic reporting to journals and conferences, based on the definition of Jedlitschka et al. (2008) and Wohlin et al. (2012).
Work products	**Input:** P1.1.1 – Educational game [Ready to use]; P4.1.1 – MEEGA+ data analysis spreadsheets [Ready to use]; P1.2.2 - Evaluation goal [Ready to use]; P2.1.1 – Evaluation context [Ready to use]; P2.2.1 – Evaluation schedule [Ready to use]; P2.4.1 – Approval of the Ethics Committee [Ready to use]; P4.2.1 – Descriptive statistics results [Ready to use]; P4.3.2 – Game quality level [Ready to use]; P4.4.1 – Evaluation results [Ready to use]; P4.5.1 – Discussion [Ready to use]
	Output: P5.1.1 – Evaluation report [Ready to use]

3.3 Definition of the work products (outputs) of the MEEGA+ Process

Phase 1. Scoping

Output: P1.1.1 – Educational game

Game's name	Indicates the game's name.
Platform	(Digital or non-digital)
Computing knowledge area	Indicates the computing knowledge area following the ACM and IEEE curriculum guidelines for undergraduate degree programs in Computer Science (ACM/IEEE-CS, 2013). The Knowledge Areas are (ACM/IEEE-CS, 2013): AL - Algorithms and Complexity AR - Architecture and Organization

	CN - Computational Science DS - Discrete Structures GV - Graphics and Visualization HCI - Human-Computer Interaction IAS - Information Assurance and Security IM - Information Management IS - Intelligent Systems NC - Networking and Communications OS - Operating Systems PBD - Platform-based Development PD - Parallel and Distributed Computing PL - Programming Languages SDF - Software Development Fundamentals SE - Software Engineering SF - Systems Fundamentals SP - Social Issues and Professional Practice
Time	Indicates the duration of a game session in minutes.
Learning objectives	Indicates the learning objectives of the selected game.

Output: P1.2.2 - Evaluation goal

Evaluation goal	Analyse the <*name of the selected game*> for the purpose of *evaluate the quality* in terms of *usability and player experience* from the ***students'*** point of view in the context of *higher computing education*.

Phase 2. Planning

Output: P2.1.1 – Evaluation context

Institution	Indicates the institution's name.
Program	Indicates the name of the program.
Course	Indicates the name of the course.
Population	Indicates the number of students enrolled in the course.

Output: P2.2.1 – Evaluation schedule

Date and hour	Indicates the date and hour that the evaluation will be conducted.
Room	Indicates the room/lab number and/or address that the evaluation will be conducted.

Output: P2.3.1 – MEEGA+ data collection instruments (Digital game)
The MEEGA+ questionnaires are available in English, Spanish, and Brazilian Portuguese at http://www.gqs.ufsc.br/meega-a-model-for-evaluating-educational-games/.

Questionnaire for quality evaluation of digital games

Game: _____

Please, help us improve the game answering the following questions. All information is collected anonymously and will be used only in a summarized way in the context of this game evaluation.

Researcher: _____
Place and date: _____

Demographic Information	
Institution:	
Program:	
Course:	
Age group:	☐ Under 18 years ☐ 18 to 28 years ☐ 29 to 39 years ☐ 40 to 50 years ☐ Over 50 years
Gender:	☐ Male ☐ Female
How often do you play digital games?	☐ Never ☐ Rarely: from time to time ☐ Monthly: at least once a month ☐ Weekly: at least once a week ☐ Daily: every day.
How often do you play non-digital games (card or board games, etc.)?	☐ Never ☐ Rarely: from time to time ☐ Monthly: at least once a month ☐ Weekly: at least once a week ☐ Daily: every day.

Please, **select an option** according to how much you agree or disagree with each statement below.

Player Experience						
Statements	**Select an option as your evaluation**					
	Strongly disagree	**Disagree**	**Neither disa-gree nor agree**	**Agree**	**Strongly agree**
The game design is attractive (interface, graphics, etc.).	☐	☐	☐	☐	☐
The text font and colors are well blended and consistent.	☐	☐	☐	☐	☐
I needed to learn a few things before I could play the game.	☐	☐	☐	☐	☐
Learning to play this game was easy for me.	☐	☐	☐	☐	☐
I think that most people would learn to play this game very quickly.	☐	☐	☐	☐	☐
I think that the game is easy to play.	☐	☐	☐	☐	☐
The game rules are clear and easy to understand.	☐	☐	☐	☐	☐
The fonts (size and style) used in the game are easy to read.	☐	☐	☐	☐	☐
The colors used in the game are meaningful.	☐	☐	☐	☐	☐
The game allows customizing the appearance (font and/or color) according to my preferences.	☐	☐	☐	☐	☐
The game prevents me from making mistakes.	☐	☐	☐	☐	☐
When I make a mistake, it is easy to recover from it quickly.	☐	☐	☐	☐	☐
When I first looked at the game, I had the impression that it would be easy for me.	☐	☐	☐	☐	☐
The contents and structure helped me to become confident that I would learn with this game.	☐	☐	☐	☐	☐
This game is appropriately challenging for me.	☐	☐	☐	☐	☐
The game provides new challenges (offers new obstacles, situations or variations) at an appropriate pace.	☐	☐	☐	☐	☐
The game does not become monotonous as it progresses (repetitive or boring tasks).	☐	☐	☐	☐	☐
Completing the game tasks gave me a satisfying feeling of accomplishment.	☐	☐	☐	☐	☐

It is due to my personal effort that I managed to advance in the game.	☐	☐	☐	☐	☐
I feel satisfied with the things that I learned from the game.	☐	☐	☐	☐	☐
I would recommend this game to my colleagues.	☐	☐	☐	☐	☐
I was able to interact with other players during the game.	☐	☐	☐	☐	☐
The game promotes cooperation and/or competition among the players.	☐	☐	☐	☐	☐
I felt good interacting with other players during the game.	☐	☐	☐	☐	☐
I had fun with the game.	☐	☐	☐	☐	☐
Something happened during the game (game elements, competition, etc.) which made me smile.	☐	☐	☐	☐	☐
There was something interesting at the beginning of the game that captured my attention.	☐	☐	☐	☐	☐
I was so involved in my gaming task that I lost track of time.	☐	☐	☐	☐	☐
I forgot about my immediate surroundings while playing this game.	☐	☐	☐	☐	☐
The game contents are relevant to my interests.	☐	☐	☐	☐	☐
It is clear to me how the contents of the game are related to the course.	☐	☐	☐	☐	☐
This game is an adequate teaching method for this course.	☐	☐	☐	☐	☐
I prefer learning with this game to learning through other ways (e.g. other teaching methods).	☐	☐	☐	☐	☐

Perceived Learning					
Statements	**Select an option as your evaluation**				
	Strongly disagree	**Disagree**	**Neither disa-gree nor agree**	**Agree**	**Strongly agree**
The game contributed to my learning in this course.	☐	☐	☐	☐	☐
The game allowed for efficient learning compared with other activities in the course	☐	☐	☐	☐	☐
This statement is repeated for each learning goal of the game. The game contributed to <*verb as level of the learning goal (cognitive, psychomotor, affective)*> <*goal/concept*>. An example in accordance with the learning goals of SCRUMIA (Gresse von Wangenheim et al., 2013): The game contributed to *recall* the concepts from *Sprint Planning*.	☐	☐	☐	☐	☐

Please list three strong aspects of the game:_____

_____

_____

Please give three suggestions to improve the game:_____

_____

_____

Any further comment?_____

_____

_____

Thanks a lot for your contribution!

Output: P2.3.1 – MEEGA+ data collection instruments (Non-digital game)
The MEEGA+ questionnaires are available in English, Spanish, and Brazilian Portuguese at http://www.gqs.ufsc.br/meega-a-model-for-evaluating-educational-games/.

Questionnaire for quality evaluation of non-digital games

Game: _____

Please, help us improve the game answering the following questions. All information is collected anonymously and will be used only in a summarized way in the context of this game evaluation.

Researcher: _____
Place and date: _____

Demographic Information	
Institution:	
Program:	
Course:	
Age group:	☐ Under 18 years ☐ 18 to 28 years ☐ 29 to 39 years ☐ 40 to 50 years ☐ Over 50 years
Gender:	☐ Male ☐ Female
How often do you play digital games?	☐ Never ☐ Rarely: from time to time ☐ Monthly: at least once a month ☐ Weekly: at least once a week ☐ Daily: every day.
How often do you play non-digital games (card or board games, etc.)?	☐ Never ☐ Rarely: from time to time ☐ Monthly: at least once a month ☐ Weekly: at least once a week ☐ Daily: every day.

Please, **select an option** according to how much you agree or disagree with each statement below.

Player Experience					
Statements	**Select an option as your evaluation**				
	Strongly disagree	**Disagree**	**Neither disa-gree nor agree**	**Agree**	**Strongly agree**
The game design is attractive (boards, cards, etc.).	☐	☐	☐	☐	☐
The text font and colors are well blended and consistent.	☐	☐	☐	☐	☐
I needed to learn a few things before I could play the game.	☐	☐	☐	☐	☐
Learning to play this game was easy for me.	☐	☐	☐	☐	☐
I think that most people would learn to play this game very quickly.	☐	☐	☐	☐	☐
I think that the game is easy to play.	☐	☐	☐	☐	☐
The game rules are clear and easy to understand.	☐	☐	☐	☐	☐
The fonts (size and style) used in the game are easy to read.	☐	☐	☐	☐	☐
The colors used in the game are meaningful.	☐	☐	☐	☐	☐
When I first looked at the game, I had the impression that it would be easy for me.	☐	☐	☐	☐	☐
The contents and structure helped me to become confident that I would learn with this game.	☐	☐	☐	☐	☐
This game is appropriately challenging for me.	☐	☐	☐	☐	☐

The game provides new challenges (offers new obstacles, situations or variations) at an appropriate pace.	☐	☐	☐	☐	☐
The game does not become monotonous as it progresses (repetitive or boring tasks).	☐	☐	☐	☐	☐
Completing the game tasks gave me a satisfying feeling of accomplishment.	☐	☐	☐	☐	☐
It is due to my personal effort that I managed to advance in the game.	☐	☐	☐	☐	☐
I feel satisfied with the things that I learned from the game.	☐	☐	☐	☐	☐
I would recommend this game to my colleagues.	☐	☐	☐	☐	☐
I was able to interact with other players during the game.	☐	☐	☐	☐	☐
The game promotes cooperation and/or competition among the players.	☐	☐	☐	☐	☐
I felt good interacting with other players during the game.	☐	☐	☐	☐	☐
I had fun with the game.	☐	☐	☐	☐	☐
Something happened during the game (game elements, competition, etc.) which made me smile.	☐	☐	☐	☐	☐
There was something interesting at the beginning of the game that captured my attention.	☐	☐	☐	☐	☐
I was so involved in my gaming task that I lost track of time.	☐	☐	☐	☐	☐
I forgot about my immediate surroundings while playing this game.	☐	☐	☐	☐	☐
The game contents are relevant to my interests.	☐	☐	☐	☐	☐
It is clear to me how the contents of the game are related to the course.	☐	☐	☐	☐	☐
This game is an adequate teaching method for this course.	☐	☐	☐	☐	☐
I prefer learning with this game to learning through other ways (e.g. other teaching methods).	☐	☐	☐	☐	☐

Perceived Learning					
Statements	**Select an option as your evaluation**				
	Strongly disagree	**Disagree**	**Neither disa-gree nor agree**	**Agree**	**Strongly agree**
The game contributed to my learning in this course.	☐	☐	☐	☐	☐
The game allowed for efficient learning compared with other activities in the course	☐	☐	☐	☐	☐
This statement is repeated for each learning goal of the game. The game contributed to <*verb as level of the learning goal (cognitive, psychomotor, affective)*> <*goal/concept*>. An example in accordance with the learning goals of SCRUMIA (Gresse von Wangenheim et al., 2013): The game contributed to *recall* the concepts from *Sprint Planning*.	☐	☐	☐	☐	☐

Please list three strong aspects of the game:_____
_____
_____

Please give three suggestions to improve the game:_____
_____
_____

Any further comment?_____
_____
_____

Thanks a lot for your contribution!

Output: P2.4.1 – Approval of the Ethics Committee (optional)

There is no a template for this output. This output is defined by the Ethics Committee of the institution that the evaluation will be conducted.

Phase 3. Execution

Output: P3.1.2 – Consent form printed (optional)

There is no a template for this output. A template for this output is typically provided by the Ethics Committee of the institution that the evaluation will be conducted.

Output: P3.2.1 – Participants' consent (optional)

A document presenting the formal consent signed by the participants, following the template provided in the Output: P3.1.2 – Consent form printed.

Output: P3.3.1 – Game executed

Duration:	Indicates the duration of the game execution
Execution:	Describes the execution of the game and any deviations from plan. If the game is a digital game, describes if technological problems occurred. If the game is a non-digital game, describes any problem with the room, for example, the tables available were small for the board used in the game.

Output: P3.4.1 – Data collected
All data collection instruments/electronic forms filled out by the participants (following the template of the output P2.3.1 – MEEGA+ model data collection instruments) and checked/validated by the researcher.

Sample size:	Number of students that participated of the evaluation and filled out the questionnaire in a valid way.

Phase 4. Analysis

Output: P4.1.1 – MEEGA+ data analysis spreadsheets
The MEEGA+ data analysis spreadsheets are available in English, Spanish, and Brazilian Portuguese at http://www.gqs.ufsc.br/meega-a-model-for-evaluating-educational-games/.

The MEEGA+ data analysis spreadsheet contains five internal spreadsheets (Data; Demographic data; Graphs, Demographic graphs; and Items).

The Data spreadsheet (Figure 14) aims to organize the data collected for a descriptive analysis. Each row of the spreadsheet represents an answer of one student for the game evaluation. The columns represent the MEEGA+ measurement instrument items. After inserting the data collected in the spreadsheet, the average, median and frequency of responses are automatically calculated.

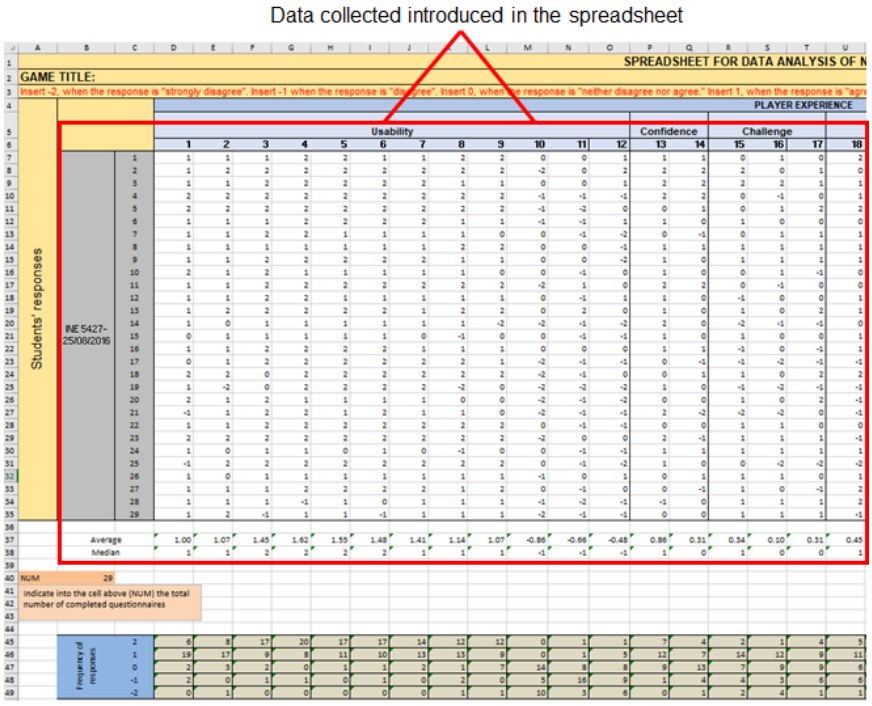

Figure 14. **Data analysis spreadsheet (evaluation data)**

The Demographic data spreadsheet (Figure 15) aims to organize the demographics data collected through the MEEGA+ measurement instrument, in order to identify the characteristics of the sample. The demographics information collected are age group, gender and how often students play digital and/or non-digital games.

	A	B	C	D	E	F	G
1				DEMOGRAPHIC INFORMATION			
2				**Age group:** 1 - Under 18 years 2 - 18 to 28 years 3 - 29 to 39 years 4- 40 to 50 years 5 - Over 50 years	**Gender:** M - Male F - Female	**How often do you play digital games?** 1 - Never 2 - Rarely: from time to time 3 - Monthly: at least once a month 4 - Weekly: at least once a week 5 - Daily: every day.	**How often do you play non-digital games (card or board games, etc.)?** 1 - Never 2 - Rarely: from time to time 3 - Monthly: at least once a month 4 - Weekly: at least once a week 5 - Daily: every day.
3	Students' responses	INE 5617 - 31/03/2017	1	2	M	4	3
4			2	2	M	4	3
5			3	2	M	4	2
6			4	2	M	5	3
7			5	2	M	5	2
8			6	2	M	5	2
9			7	2	M	5	2
10			8	2	F	4	2
11			9	3	F	2	3
12			10	3	M	2	2
13			11	2	F	3	2
14			12	3	M	2	2
15			13	2	F	3	2
16			14	2	M	4	3
17			15	2	M	5	3
18			16	2	M	2	2
19			17	2	M	4	2
20			18	2	M	4	3
21			19	2	M	3	4
22			20	2	M	5	4
23			21	2	F	5	3
24			22	2	M	5	4

Figure 15. **Data analysis spreadsheet (demographic data)**

The Items spreadsheet (Figure 16) is a standardized spreadsheet that contains the description of the MEEGA+ measurement instrument items used in the graphs. The customized items related to the learning goals of the evaluated game should be inserted in this spreadsheet in the corresponding dimension (Learning goals) to be used in the graphs.

	A	B	C	D	E
1	**Quality factor**	**Dimension**		**No. Item**	**Item Description**
2	Usability	Aesthetics		1	The game design is attractive (interface, graphics, etc.).
3				2	The text font and colors are well blended and consistent.
4		Learnability		3	I needed to learn a few things before I could play the game.
5				4	Learning to play this game was easy for me.
6				5	I think that most people would learn to play this game very quickly.
7		Operability		6	I think that the game is easy to play.
8				7	The game rules are clear and easy to understand.
9		Accessibility		8	The fonts (size and style) used in the game are easy to read.
10				9	The colors used in the game are meaningful.
11				10	The game allows customizing the appearance (font and/or color) according to my preferences.
12	Player experience	User error protection		11	The game prevents me from making mistakes.
13				12	When I make a mistake it is easy to recover from it quickly.
14		Confidence		13	When I first looked at the game, I had the impression that it would be easy for me.
15				14	The contents and structure helped me to become confident that I would learn with this game.
16		Challenge		15	This game is appropriately challenging for me.
17				16	The game provides new challenges (offers new obstacles, situations or variations) at an appropriate pace.
18				17	The game does not become monotonous as it progresses (repetitive or boring tasks).
19		Satisfaction		18	Completing the game tasks gave me a satisfying feeling of accomplishment.
20				19	It is due to my personal effort that I managed to advance in the game.
21				20	I feel satisfied with the things that I learned from the game.
22				21	I would recommend this game to my colleagues.
23		Social Interaction		22	I was able to interact with other players during the game.
24				23	The game promotes cooperation and/or competition among the players.
25				24	I felt good interacting with other players during the game.
26		Fun		25	I had fun with the game.
27				26	Something happened during the game (game elements, competition, etc.) which made me smile.
28		Focused Attention		27	There was something interesting at the beginning of the game that captured my attention.
29				28	I was so involved in my gaming task that I lost track of time.
30				29	I forgot about my immediate surroundings while playing this game.
31		Relevance		30	The game contents are relevant to my interests.
32				31	It is clear to me how the contents of the game are related to the course.
33				32	This game is an adequate teaching method for this course.
34				33	I prefer learning with this game to learning through other ways (e.g. other teaching methods).
35		Perceived Learning		34	The game contributed to my learning in this course.
36				35	The game allowed for efficient learning compared with other activities in the course
37		Learning goals		36	The game contributed for review the knowledge of scope planning

Figure 16. Items spreadsheet

Output: P4.2.1 – Descriptive statistics results

Frequency of responses, average and median for each measurement instrument item (Figure 17).

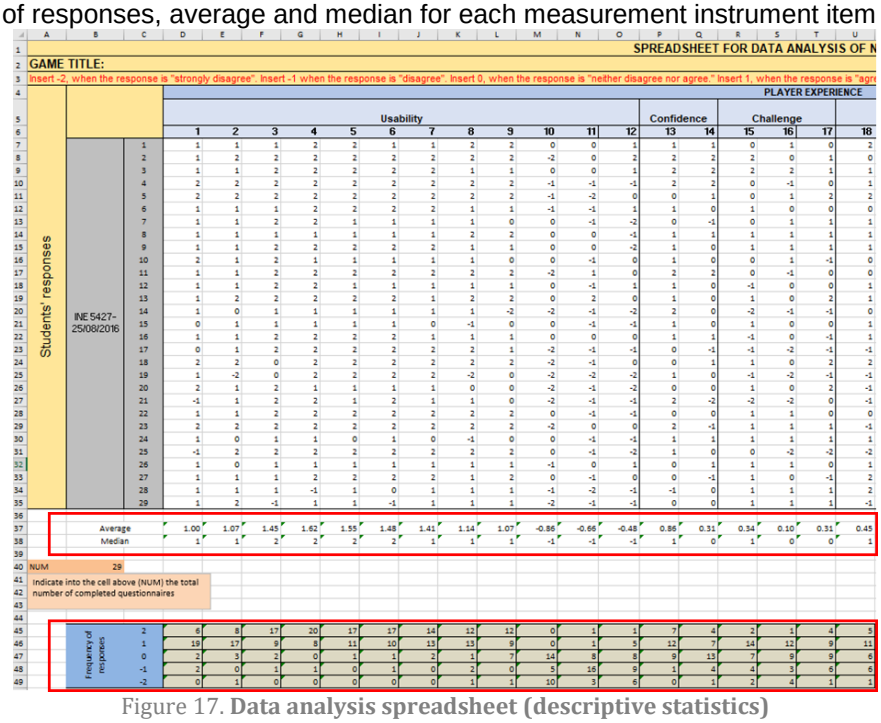

Figure 17. **Data analysis spreadsheet (descriptive statistics)**

Graphics of frequency are presented in accordance with the quality factors and dimensions of the MEEGA+ model (Figure 18).

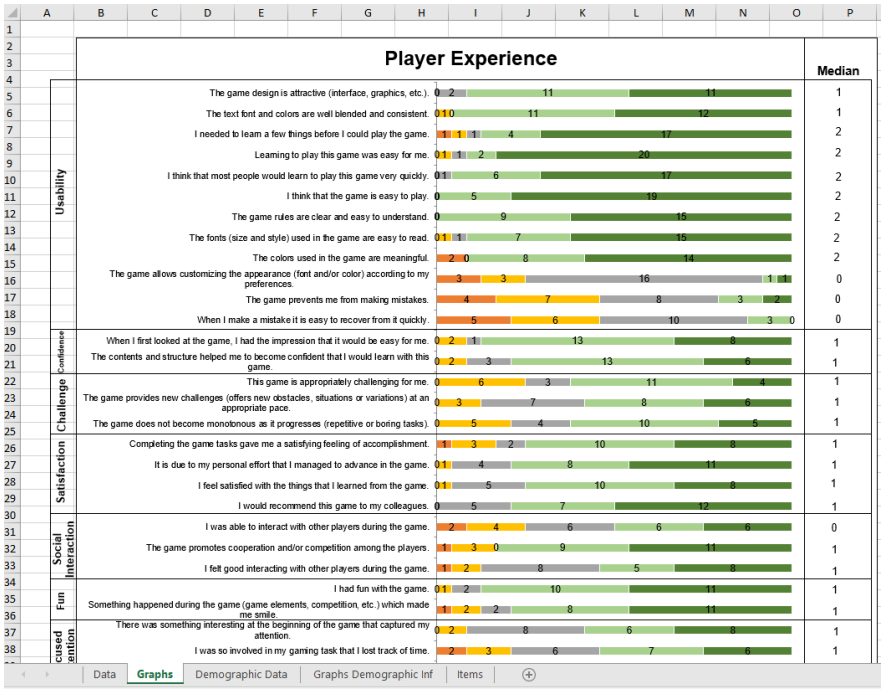

Figure 18. **Frequency graphs**

Graphics presenting the demographic information are generated and presented in the Demographic graphs spreadsheet (Figure 19).

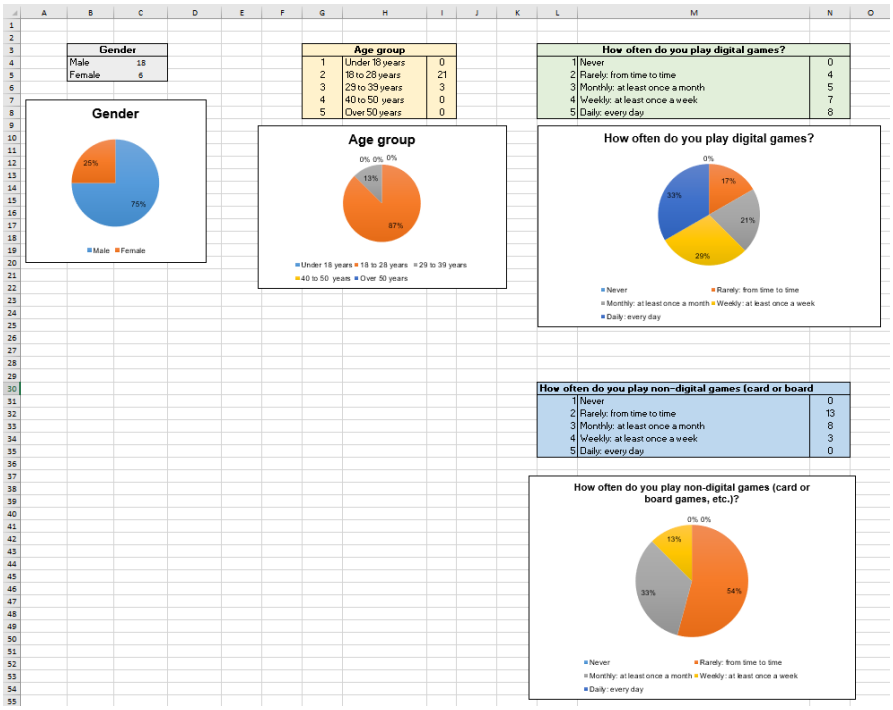

Figure 19. **Demographic information graphs**

Output: P4.3.2 – Game quality level

<table>
<tr><td>Game quality level:</td><td>Describes the characteristics of the evaluated game based on their quality level, as described in Table 11.
Table 11. Game quality levels
<table>
<tr><th>Quality level</th><th>Level description</th></tr>
<tr><td>Low quality
(θ < 42.5)</td><td>At this level, the game rarely provides social interaction and hardly ever produces moments of fun among the players. The game does not capture the student's focused attention, does not arouse the confidence that he/she will learn from the game, nor does it produce feelings of satisfaction. The game rarely presents challenges, has monotonous tasks and does not contribute to student learning. Although a game at this level has a low relevance to the students' interests, a student recognizes that the game's content is related to the course. In terms of usability, a game at this level sometimes exhibits operability features, which may have some clear rules and be easy to play.</td></tr>
<tr><td>Good quality
(42.5 <= θ < 65)</td><td>At this level, the game sometimes presents challenging activities, offering new challenges for students. It provides moderately focused attention to the players, although students do not forget about their surroundings. Sometimes the game also provides feelings of confidence and satisfaction in the players. Frequently the game presents moments of social interaction and fun among the players. Often the game is considered relevant to the students' interests and, usually, the students recognize that the game's content is related to the course. Frequently the game contributes efficiently to student learning. In terms of usability, the game usually has the clear rules and is easy to play, although, usually does not present a fully attractive design.</td></tr>
<tr><td>Excellent quality
(θ >= 65)</td><td>At this level, the game is challenging for students and has no monotonous activities. It is highly relevant to students' interests and provides excellent focused attention, satisfaction, fun, and social interaction. It allows the student to be confident that he/she will learn from the game and contribute to and efficient student learning. In terms of usability, the game presents excellent operability and learnability, that is, it has clear rules and is easy to learn to play. Even so, a game at this level may present improvements in terms of aesthetics, not presenting a fully attractive design.</td></tr>
</table></td></tr>
</table>

Output: P4.4.1 – Evaluation results

Analysis question 1: Does the <*name of the evaluated game*> has a good usability?	Analyses the main results for each dimension of the quality factor usability, answering the defined analysis question.
Analysis question 2: Does the <*name of the evaluated game*> provides a positive player experience?	Analyses the main results for each dimension of the quality factor player experience, answering the defined analysis question.
Analysis question 3: How old are the students that compose the sample of the study?	Analyses the frequency of responses to each age group.
Analysis question 4: What is the gender of the students that compose the sample of the study?	Analyses the frequency of responses to each gender.
Analysis question 5: What is the frequency that the students play digital and/or non-digital games?	Analyses the frequency of each response category indicating how often the students play digital and/or non-digital games.

Output: P4.5.1 – Discussion

Discussion:	Discussion of findings and, optionally, compared them with similar researches.
Threats to validity:	Indicates the limitations of the study and its threats to validity, as well as mitigation strategies adopted in order to minimize the impact in the study.

Phase 5. Presentation

Output: P5.1.1 – Evaluation report

Section	**Contents/Scope**	
Title	Evaluation of the <*object of study*> for the <*computing knowledge area*> education in <*evaluation context*>. E.g.: Evaluation of the Kahoot! PMQuiz for the Project Management education in Computing courses.	
Authorship	Names, contacts and affiliations.	
Abstract	Summarizes the paper under headings of background or context, objectives, method, main results, and conclusions.	
Introduction	Problem Statement	Indicates what the problem is; where is occurs, and who observe it.
	Research Objective	Defines the evaluation using the formalized style used in the MEEGA+ model.
	Context	Indicates environmental factors such as institution, course, and participants involved in the evaluation.
Related Work	How this research relates to existing research (studies)?	
Research Method	Reports the methods used in the research, such as the MEEGA+ method (Petri et al., 2018), case studies (Yin, 2017; Wohlin et al., 2012), GQM (Basili et al., 1994).	
Evaluation Planning	Object of study	Indicates the game selected fort the evaluation.
	Evaluation goal	Presents the defined evaluation goal and the analysis questions following the MEEGA+ model. Analyse the <*name of the selected game*> for the purpose of *evaluate the quality* in terms of *usability and player experience* from the ***students'*** point of view in the context of *higher computing education*.
	Analysis questions	AQ1: Does the <*name of the evaluated game*> has a good usability? AQ2: Does the <*name of the evaluated game*> provides a positive player experience? AQ3: How old are the students that compose the sample of the study? AQ4: What is the gender of the students that compose the sample of the study? AQ5: What is the frequency that the students play digital and/or non-digital games?
	Context details	Indicates the place that the evaluation took place, such as institution and course.
	Research design	Indicates the research design applied, following the definition of the MEEGA+ model. Case study design (one-shot post-test only).
	Schedule	Indicates the schedule of the evaluation such as date and time.
	Number of	Indicates the number of the approval provided by the Ethics

	the Ethics Committee approval	Committee (if necessary).
Execution	Sample	Description of the sample characteristics (demographic information).
	Preparation	What has been done to prepare the execution of the evaluation (i.e., schedule, materials)?
	Game applied	Indicates how game application took place and any deviations from plan.
	Data collection performed	How data collection took place and any deviations from plan.
Analysis	Answer the analysis questions	Summarizes the data collected and describes how it was analysed and answers each of the analysis questions defined.
	Game quality level	Indicates the quality level of the evaluated game, obtained from the MEEGA+ scale.
Discussion	Evaluation of results	Interprets and explains the findings from the Analysis section.
	Threats to validity	Discusses the main threats to validity and mitigation strategies applied.
Conclusions and Future Work	Summary	Provides a concise summary of the research objective and evaluation execution.
	Findings	Identifies the most important results of the study.
	Improvement opportunities	Suggestions for other studies to further investigate.
Acknowledgements	Identifies any sponsors, participants, and contributors who do not fulfil authorship criteria.	
References	Lists all cited literature in the format requested by the publisher.	
Appendices	Includes supplementary data and/or detailed analyses which might help others to use the results.	

4. Conclusions

In this technical report, we presented the MEEGA+ method, an evolution of an evaluation model of educational games used as instructional strategy for computing education, improving the initial model proposed by Savi et al. (2011). The MEEGA+ method aims to provide a systematic support for the evaluation of games for computing education, focusing on the quality evaluation of educational games (including digital as well as non-digital games) in terms of usability and player experience. It is composed of an evaluation model (MEEGA+ model) defining quality aspects to evaluate a game, and a process (MEEGA+ process) guiding the conduction of the game evaluation.

As next steps, we plan to continue conducting case studies evaluating games (digital and non-digital) for computing education using the MEEGA+ method in order to conduct a statistical analysis to confirm the decomposition of the MEEGA+ factors and dimensions. In addition, we plan evaluate the quality of the MEEGA+ method from the expert's perspective.

Acknowledgments

We would like to thank all the students and instructors that accepted participate in the applications of the games using the MEEGA+ method.

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.

This work was partially conducted during a visiting scholar period at University of Cádiz, sponsored by the CAPES (*Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*), Foundation within the Ministry of Education, Brazil (grant n. 88881.131485/2016-01).

References

Abt, C. C. (2002). *Serious Games*. Lanhan: University Press of America.

ACM/IEEE-CS. (2013). *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 Nov. 2017.

Acuña, S. T.; Antonio, A.; Ferré, X.; López, M.; Maté, L. (2000). The Software process: modeling, evaluation and improvement. *Handbook of Software Engineering and Knowledge Engineering*. World Scientific Publishing Company.

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

Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. Longman.

Andrade, D. F., Tavares, H. R., Valle, R. C. (2000). Teoria de Resposta ao Item: conceitos e aplicações. ABE — Associação Brasileira de Estatística, 4º SINAPE.

Andrade, H. & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory into Practice*, 48, 12-19.

Basili, V. R., Caldiera, G., & Rombach, H. D. (1994). Goal, Question Metric Paradigm. In J. J. Marciniak, *Encyclopedia of Software Engineering*, (pp. 528-532). New York: Wiley-Interscience.

Battistella, P. & Gresse von Wangenheim, C. (2016). Games for teaching computing in higher education – a systematic review. *IEEE Technology and Engineering Education Journal*, 9(1), 8-30.

Beecham, S., Hall, T., Britton, C., Cottee, M., & Rainer, A. (2005). Using an Expert Panel to Validate a Requirements Process Improvement Model. *Journal of Systems and Software*, 76(3), 251-275.

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

Boyle, E. A., Connolly, T. M., & Hainey, T. (2011). The role of psychology in understanding the impact of computer games. *Entertainment Computing*, 2, 69–74.

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., Ribeiro, C., & Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education*, 94, 178-192.

Branch, R. M. (2010). *Instructional Design: The ADDIE Approach.* New York: Springer New York Dordrecht Heidelberg London.

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability Evaluation in Industry*, 189(194), 4-7.

Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan (Ed.). *The SAGE handbook of research on classroom assessment* (pp. 367-393). Thousand Oaks: Sage Publications.

Brown, G., Andrade, H., & Chen, F. (2015). Accuracy in student self-assessment: Directions and cautions for research. *Assessment in Education Principles Policy and Practice*, 22(4), 1-26.

Budgen, D., Turner, M., Brereton, P., & Kitchenham, B. (2008). Using mapping studies in Software Engineering. *Proc. of the 20th Annual Workshop of Psychology of Programming Interest Group*, (pp. 195-204). Lancaster University, Lancaster, USA.

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

Calderón, A., Ruiz M., & O'Connor, R. (2018). A multivocal literature review on serious games for software process standards education. *Computer Standards & Interfaces*, 57, 36-48.

Çiftci, S. (2018). Trends of Serious Games Research from 2007 to 2017: A Bibliometric Analysis. *Journal of Education and Training Studies*, 6(2), 18-27.

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

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales*. International Journal of Market Research,* 50(1), 61-77.

DeVellis, R. F. (2016). *Scale development: theory and applications* (4th. ed.). Thousand Oaks: SAGE Publications.

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

Dolan, E. L. & Collins, J. P. (2015). We must teach more effectively: here are four ways to get started. *Molecular Biology of the Cell*, 26(12), 2151-2155.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proc. of the National Academy of Sciences of the United States of America*, 111(23), 8410-8415.

Fu, F., Su, R., & Yu, S. (2009). EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education*, 52(1), 101-112.

Gámez, E. H. (2009). *On the Core Elements of the Experience of Playing Video Games* (Dissertation). UCL Interaction Centre, Department of Computer Science, London.

Gibson, B. & Bell, T. (2013). Evaluation of games for teaching computer science. *Proc. of the 8th Workshop in Primary and Secondary Computing Education*, (pp. 51-60). ACM, New York, USA.

Gomes, M. N. (2016). Desenvolvimento de uma Atividade Gamificada Voltada ao Turismo de Caxias Do Sul. Undergraduate Thesis (Computer Science Program), Universidade de Caxias do Sul (in Portuguese).

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

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

Hambleton, R. K., Swaminathan, H. and Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park: Sage Publications.

Herpich, F., Nunes, F. B., Voss, G. B., Sindeaux, P., Tarouco, L. M. R., & Lima, J. V. (2017). Realidade Aumentada em Geografia: uma atividade de orientação no ensino fundamental. *Revista Novas Tecnologias na Educação*, 15(2), 1–10 (in Portuguese).

IEEE. Standard Glossary of Software Engineering Terminology. Std 610.12-1990(R2002), 2002.

IEEE. Systems and Software Engineering – Vocabulary. ISO/IEC/IEEE 24765:2010(E), pp.1-418, 2010.

International Standard Organization (ISO). (2014). ISO/IEC 25010: Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – System and software quality models, Technical Report.

Izquierdo, M. C. & Pedrero, E. F. (2014). Estimating the reliability coefficient of tests in presence of missing values. *Psicothema*, 26 (4), 516-523.

Jedlitschka A., Ciolkowski M., Pfahl D. (2008) *Reporting Experiments in Software Engineering*. In: Shull F., Singer J., Sjøberg D.I.K. (eds) Guide to Advanced Empirical Software Engineering. London: Springer.

Johnson, R. B. & Christensen, L. (2014). *Educational Research – Quantitative, Qualitative, and Mixed Approaches* (5th ed.). Thousand Oaks: Sage Publications, Inc.

Kasunic, M. (2005). Designing an effective survey. Handbook CMU/SEI-2005-HB-004, Software Engineering Institute. Pittsburgh: Carnegie Mellon University.

Keller, J. (1987). Development and Use of the ARCS Model of motivational Design. *Journal of Instructional Development*, 10(3), 2-10.

Kitchenham, B. (2010). Systematic literature reviews in software engineering – A tertiary study. *Information and Software Technology*, 52(1), 792-805.

Kordaki, M. & Gousiou, A. (2016). Computer Card Games in Computer Science Education: A 10-Year Review. *Journal of Educational Technology & Society*, 19(4), 11-21.

Kordaki, M. & Gousiou, A. (2017). Digital card games in education: A ten year systematic review. *Computers & Education*, 109, 122-161.

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

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1973). *Taxonomy of Educational Objectives, the Classification of Educational Goals*. Handbook II: Affective Domain. Philadelphia: David McKay Company, Inc.

Malhotra, N. K., & Birks, D. F. (2008). *Marketing Research: An Applied Approach* (3rd ed.). Philadelphia: Trans-Atlantic Publications.

Matook, S., Indulska, M. (2009). Improving the Quality of Process Reference Models: A Quality Function Deployment-Based Approach. *Decision Support Systems*, v. 47.

Mohamed, H. & Jaafar, A. (2010). Development and Potential Analysis of Heuristic Evaluation for Educational Computer Game (PHEG). *Proc. of the 5th Int. Conf. on Computer Sciences and Convergence Information Technology*, (pp. 222-227). Seoul, South Korea.

Münch, J., Armbrust, O., Kowalczyk, M., & Soto, M. (2012). *Software Process Definition and Management.* New York: Springer-Verlag Berlin Heidelberg.

O'Brien, H. L. & Toms, E. G. (2010). The Development and Evaluation of a Survey to Measure User Engagement. *Journal of the American Society for Information Science and Technology*, 61(1), 50–69.

Parsons, P. (2011). Preparing computer science graduates for the 21st Century. *Teaching Innovation Projects*, 1(1), article 8.

Pasquali, L., Primi, R. (2003). *Fundamentos da Teoria da Resposta ao Item –TRI*. Avaliação Psicológica, 2(2), 99-110.

Pereira, C. X., Hoffmann, Z., Castro, C. P., Santos, G. S., Aires, T. A., & Francisco, R. E. (2017). CHEMIS3: A game for learning chemical concepts through elements of nature. *Proc of the 12th Iberian Conf. on Information Systems and Technologies*, (pp. 1-5). Lisbon, Portugal.

Petri, G. & Gresse von Wangenheim, C. (2016). How to evaluate educational games: a systematic literature review. *Journal of Universal Computers Science*, 22(7), 992-1021.

Petri, G. & Gresse von Wangenheim, C. (2017). How games for computing education are evaluated? A systematic literature review. *Computers & Education*, 107, 68-90.

Petri, G., Gresse von Wangenheim, C., & Borgatto, A. F. (2017). A Large-scale Evaluation of a Model for the Evaluation of Games for Teaching Software Engineering. *Proc. of the IEEE/ACM 39th Int.*

Conf. on Software Engineering: Software Engineering Education and Training Track, (pp. 180-189). Buenos Aires/Argentina.

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

Petri, G., Gresse Von Wangenheim, C., & Borgatto, A. F. (2018b). Design and Evaluation of a Model for the Evaluation of Games for Computing Education. *Journal of Informatics in Education* (under review).

Petri, G., Gresse Von Wangenheim, C., & Borgatto, A. F. (2018c). *Classifying the quality level of educational games: the MEEGA+ game scale*. Journal of Informatics in Education (under review).

Poels, K., Kort, Y. D., & Ijsselsteijn, W. (2007). It is always a lot of fun!: exploring dimensions of digital game experience using focus group methodology. *Proc. of Conf. on Future Play*, (pp. 83-89). Toronto, Canada.

Prensky, M. (2007). *Digital Game-Based Learning*. New York: Paragon House.

Ritterfeld, U., Cody, M., & Vorderer, P. (Eds.). (2010). *Serious Games*. New York: Routledge.

Rittgen, P. (2010). Quality And Perceived Usefulness of Process Models. *Proc. of the Symposium on Applied Computing*. Nova York: ACM, 2010, p. 65-72.

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research & Evaluation*, 11(10), 1-13.

Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores*. Psychometric Monograph. Educational Testing Service. Princeton, New Jersey.

Santos, A. L., Souza, M. R. A., Figueiredo, E., & Dayrell, Marcella. (2018). Game Elements for Learning Programming: A Mapping Study. *Proc. of the 10th Int. Conf. on Computer Supported Education*, (pp. 89-101). Funchal, Portugal.

Savi, R., Gresse von Wangenheim, C., & Borgatto, A. F. (2011). A model for the evaluation of educational games for teaching software engineering. *Proc. of the 25th Brazilian Symposium on Software Engineering*, (pp. 194-203). São Paulo, Brazil (in Portuguese).

Schanzenbach, D. W. (2012). Limitations of Experiments in Education Research. *Education Finance and Policy*, 7(2), 219-232.

Sharma, R., Jain, A., Gupta, N., Garg, S., Batta, M., & Dhir, S. K. (2016). Impact of self-assessment by students on their learning. *International Journal of Applied and Basic Medical Research*, 6(3), 226–229.

Silva, J. P., Carvalho, F. C., Silvestre, A., Luiz, M. F., Alencar, R. L., & Silveira, I. F. (2017a). Forca dos Sinônimos: Um Jogo da Forca Multiplayer para o ensino de Tesauros em Língua Portuguesa como Ferramenta de busca na Web. *Prof. of the 16th Brazilian Symposium on Games & Digital Entertainment*, (pp. 1128 – 1131). Curitiba/PR, Brazil (in Portuguese).

Silva, J. P., Jorge, A. A., Gaia, C. C., & Costa, G. C. (2017b). Quais as chances? Um jogo de dados e cartas para o ensino do cálculo de probabilidades. *8º Congresso de Inovação, Ciência e Tecnologia do IFSP,* (pp. 1-3). Cubatão/SP, Brasil (in Portuguese).

Simpson, E. J. (1972). *The classification of educational objectives, psychomotor domain.* Washington: Gryphon House.

Sindre, G. & Moody, D. (2003). Evaluating the Effectiveness of Learning Interventions: An Information Systems Case Study. *Proc. of the 11th European Conf. on Information Systems*, Paper 80. Naples, Italy.

Siniscalco, M. T., Auriat, N. (2005). Questionnaire design: Quantitative research methods in educational planning. Paris: UNESCO International Institute for Educational Planning.

Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). Self-Assessment of Knowledge: A Cognitive Learning or Affective Measure? *Academy of Management Learning & Education*, 9(2), 169-191.

Sweetser, P. & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment*, 3(3), 1-24.

Tahir, R. & Wangmar, A. I. (2017). State of the art in Game Based Learning: Dimensions for Evaluating Educational Games. *Proc. of the European Conference on Games Based Learning*, (pp. 641-650). Graz, Austria.

Takatalo, J., Häkkinen, J., Kaistinen, J., & Nyman, G. (2010). Presence, Involvement, and Flow in Digital Games. In: Bernhaupt, R. (Ed.). *Evaluating User Experience in Games: Concepts and Methods*, (pp. 23-46). London: Springer.

Thomas, G., Martin, D., & Pleasants, K. (2011). Using self- and peer-assessment to enhance students' future-learning in higher education. *Journal of University Teaching & Learning Practice*, 8(1), 1-17.

Trochim, W. M., & Donnelly, J. P. (2008). *Research methods knowledge base* (3rd ed.). Mason: Atomic Dog Publishing.

Tullis, T. & Albert, W. (2008). *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. Burlington: Morgan Kaufmann.

Weske, M. (2012). *Business Process Management: Concepts, Languages, Architectures (*2nd ed.). Springer.

Wiebe, E. N., Lamb, A., Hardy, M., & Sharek, D. (2014). Measuring engagement in video game-based environments: Investigation of the User Engagement Scale. *Computers in Human Behavior*, 32, 123-132.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). *Experimentation in Software Engineering*. New York: Springer-Verlag Berlin Heidelberg.

Yin, R. K. (2017). *Case study research: design and methods* (6th ed.). Thousand Oaks: Sage Publications, Inc.

Zaibon, S. B. & Shiratuddin, N. (2010). Heuristics Evaluation Strategy for Mobile Game-Based Learning. *Proc of the 6th IEEE Int. Conf. on Wireless, Mobile and Ubiquitous Technologies in Education*, (pp. 127-131). Kaohsiung, Taiwan.

Zaibon, S. B. (2015). User Testing on Game Usability, Mobility, Playability, and Learning Content of Mobile Game-Based Learning. *Journal Teknologi*, 77(29), 131-139.