



INCoD

Instituto Nacional para Convergência Digital



UNIVERSIDADE FEDERAL
DE SANTA CATARINA

ISSN 2236-5281

Technical Report

INCoD/GQS.06.2018.E

MEEGA+KIDS: A Model for the Evaluation of Educational Games for Computing Education in Secondary School (Draft Version)

Authors:

Christiane Gresse von Wangenheim

Giani Petri

Adriano Ferreti Borgatto

Version 1.0

Status: Final

Distribution: External

AUGUST 2018

© 2018 **INCoD** – Brazilian Institute for Digital Convergence

All rights reserved and protected under Brazilian Law No. 9.610 from 19/02/1998. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise.



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+, a model to evaluate the quality of educational games used as instructional strategy for computing education, improving the initial version of MEEGA (Model for the Evaluation of Educational Games). MEEGA+ provides systematic support to evaluate the perceived quality of educational games in terms of player experience and perceived learning. The MEEGA+ model has been systematically developed by decomposing evaluation goals into measures and defining a standardized measurement instrument in form of a self-assessment questionnaire. Observing the need to evaluate also the quality of games used in computing education in middle and high school as part of an increasing tendency to popularize computing, we customized the measurement instrument to this specific target audience through a participatory design approach. As a result, this report presents a preliminary version of the questionnaire MEEGA+KIDS (in English and Brazilian Portuguese). Currently we are running an evaluation study in order to analyze the instrument's reliability and validity. We expect that the MEEGA+KIDS model provides game creators, instructors and researchers with a measurement instrument in order to evaluate the quality of educational games in secondary school and, thus, contributes to their improvement and effective and efficient adoption in practice.

1. Introduction

Teaching computing through summer camps, clubs or in family workshops is a worldwide trend (Gresse von Wangenheim & von Wangenheim, 2014). There are several initiatives to teach computing such as Code.org (<http://www.code.org>), Code.club (<https://www.codeclubworld.org>), *Computação na Escola* (<http://www.computacaonaescola.ufsc.br>), among others. These initiatives are expected to contribute to the popularization of computing competencies as well as the awareness and interest of the students towards computing (Guzdial et al., 2014; Garneli et al., 2015).

Taking into consideration the growing number of alternative instructional units (IUs) for teaching computing, it is important to obtain evidence on the expected benefits as a basis for their systematic selection, adoption and improvement (Decker et al., 2016). Following Guzdial (2004), a main contribution to this knowledge area is not necessarily the development of new programming environments or instructional units, but to find out how to study the existing ones. A more precise understanding of the results of using these instructional units would make it possible to know whether they contribute, in fact, positively to the achievement of the learning goals and compensate the cost involved in their adoption. However, although there is evidence that existing IUs can improve the teaching and learning process in middle school being used more widely in schools worldwide, there is little research on the analysis of the contribution that these IUs can bring to education (Decker et al., 2016).

Currently, the evaluation of the quality of IUs is limited or even, sometimes, non-existent (Decker et al., 2016; Garneli et al., 2015). In many cases, a decision about the use of IUs is based on assumptions of their effectiveness (Gross and Powers, 2005; Wilson et al., 2010). On the other hand, some studies focus on specific quality factors only, such as learning improvement (Gross and Powers, 2005; Kalelioğlu and Gülbahar, 2014). Other studies focus on the effectiveness of visual block-based programming languages (Weintrop and Wilensky, 2015; Grover et al., 2014; Perdikuri, 2014). However, students' perceptions and intentions are also determining factors for successful learning (Giannakos et al., 2013). Yet, few evaluations take into consideration aspects such as motivation and the students' experience during the instructional unit (Craig and Horton, 2009; Giannakos et al., 2014), or students' attitudes toward technology acceptance (Giannakos et al., 2013). In addition, studies that measure students' attitude toward computing are rather designed for higher education and seem to be outdated in the current context of teaching computing in schools (Garland and Noyes, 2008).

Especially with regard to game-based learning (Abt, 2002), there seems to be a lack of evaluation models focusing on this educational stage. Currently, there are very few approaches focusing on higher education, 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 et al., 2018). A prominent evaluation model is the MEEGA+ model (Petri et al., 2017a)(Petri et al., 2018b), an evolution of a model to evaluate the quality of educational games used as instructional strategy for computing education initially proposed by Savi et al. (2011) (Petri et al., 2018b). MEEGA+ provides systematic support to evaluate the perceived quality of educational games in terms of player experience and perceived learning. The MEEGA+ model has been systematically developed by decomposing evaluation goals into measures and defining a standardized measurement instrument in form of a self-assessment questionnaire. Furthermore, the MEEGA+ method (Petri et al., 2018a) provides 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.

Observing the need to evaluate also the quality of games used in computing education in secondary schools as part of an increasing tendency to popularize computing, we customized the measurement instrument to this specific target audience. Based on the MEEGA+ model, the measurement instrument has been adapted by revising the wording in accordance to the specific target audience adopting a participatory design approach. As a result this report presents a preliminary version of the questionnaire MEEGA+KIDS (in English and Brazilian Portuguese). Currently, we are conducting an evaluation study in order to analyze the instrument's reliability and validity. We expect that the MEEGA+KIDS model provides game creators, instructors and researchers with a measurement instrument in order to evaluate the quality of educational games in middle and high school and, thus, contribute to their improvement and effective and efficient adoption in practice.

2. Research Method

The MEEGA+KIDS model has been developed based on the MEEGA+ method (Petri et al., 2018) that has been developed systematically using a multi-method research as shown in Figure 1.

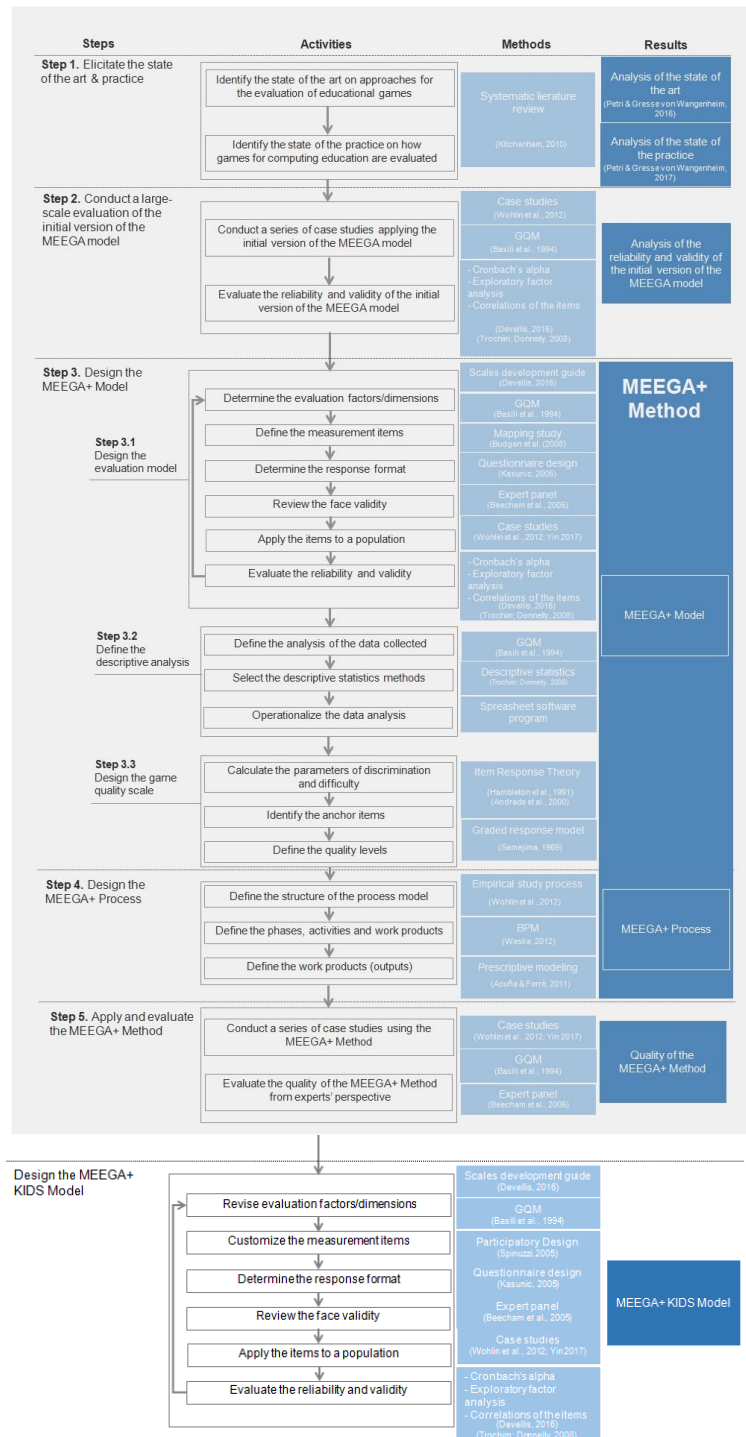


Figure 1. Research method

Based on a context analysis of the expected application of the model in computing education in secondary school, we revised the analysis questions and quality dimensions. We then customized the measurement instrument of the MEEGA+ model (Petri et al., 2018b) by adopting a participatory design approach, involving representatives of the target audience in the adaptation of the measurement items of the self-assessment questionnaire. As a result, we created the MEEGA+KIDS model (draft version) presented in this technical report.

Currently, we are conducting an evaluation study of the MEEGA+KIDS model through a series of case studies applying the model in computing education in secondary schools after the application of a game. The aim of the evaluation study is to analyze the reliability and validity of the MEEGA+KIDS model.

3. MEEGA+KIDS: A Model for the Evaluation of Games for Computing Education in Secondary School

This section presents the MEEGA+KIDS model for the quality evaluation of games used as an instructional strategy for computing education in secondary school customizing MEEGA+.

The objective of the MEEGA+KIDS 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 in secondary school.

Following the MEEGA+ model (Petri et al., 2018b), the MEEGA+KIDS model is decomposed into two quality factors and their dimensions (Figure 2). 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).

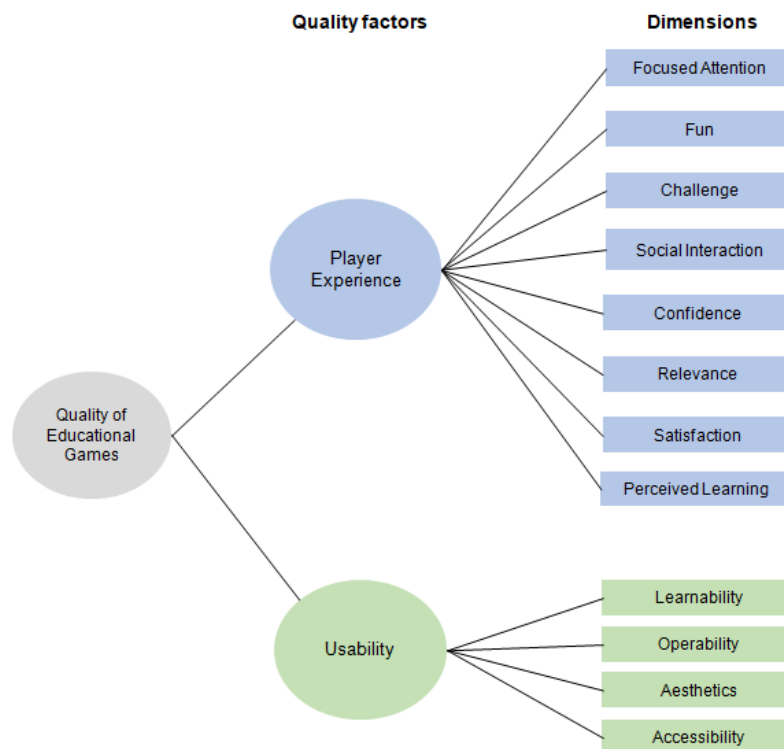


Figure 2. Decomposition of the MEEGA+KIDS model

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	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 appropriate 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+KIDS measurement instrument. Customizing the MEEGA+ measurement instrument by adopting a participatory design methodology, the MEEGA+KIDS measurement instrument has been proposed as presented in Table 2. As part of the customization some items have been excluded as considered being less important and partly covered by other items. We also changed the wording in some cases using a language more closely to the one used by the target audience to facilitate understanding.

Table 2. MEEGA+ measurement instrument items and their references

MEEGA+ (Petri et al., 2018b)			MEEGA+KIDS	
Quality factor	Dimension	Description	English draft version	Brazilian Portuguese draft version
Usability	Aesthetics (ISO/IEC, 2014)	The game design is attractive (interface, graphics, cards, boards, etc.).	The game design is attractive (game board, cards, etc.).	O design do jogo é atraente (tabuleiro, cartas, etc.).
		The text font and colors are well blended and consistent.	The font and colors of the game match.	As cores e fontes do material do jogo combinam.
	Learnability (ISO/IEC, 2014)	I needed to learn a few things before I could play the game.	--	--
		Learning to play this game was easy for me.	Learning to play this game was easy for me.	Aprender a jogar este jogo foi fácil para mim.
		I think that most people would learn to play this game very quickly.	--	--
	Operability (ISO/IEC, 2014)	I think that the game is easy to play.	I think that the game is easy to play.	Eu considero que o jogo é fácil de jogar.
		The game rules are clear and easy to understand.	The game rules are clear and easy to understand.	As regras do jogo são claras e compreensíveis.
	Accessibility (ISO/IEC, 2014)	The fonts (size and style) used in the game are easy to read.	The size and style of fonts used in the game are easy to read.	O tamanho e estilo de letras utilizadas no jogo são legíveis.
		The colors used in the game are meaningful.	The colors used in the game are meaningful.	As cores utilizadas no jogo são compreensíveis.
		The contents and structure helped me to become confident that I would learn with this game.	The organization of the content helped me to become confident that I would learn with this game.	A organização do conteúdo me ajudou a estar confiante de que eu iria aprender com este jogo.
	Challenge (Sweetser & Wyeth, 2005; Savi et al., 2011)	This game is appropriately challenging for me.	This game is appropriately challenging for me.	Este jogo é desafiador suficiente para mim.
		The game provides new challenges (offers new obstacles, situations or variations) at an appropriate pace.	The game provides new challenges (offers new obstacles, situations or variations) at an appropriate pace.	O jogo oferece novos desafios (novos obstáculos, situações ou variações) com um ritmo adequado.
		The game does not become monotonous as it progresses (repetitive or boring tasks).	The game does not become monotonous as it progresses (repetitive or boring tasks).	O jogo não se torna monótono nas suas tarefas (repetitivo ou com tarefas chatas).
	Satisfaction (Keller, 1987; Savi et al., 2011)	Completing the game tasks gave me a satisfying feeling of accomplishment.	Completing the game tasks gave me a satisfying feeling of accomplishment.	Completar as tarefas do jogo me deu um sentimento de satisfação.
		It is due to my personal effort that I managed to advance in the game.	It is due to my personal effort that I managed to advance in the game.	É devido ao meu esforço pessoal que eu consigo avançar no jogo.
		I feel satisfied with the things that I learned from the game.	I feel satisfied with the things that I learned from the game.	Me sinto satisfeito com as coisas que aprendi no jogo.

		I would recommend this game to my colleagues.	I would recommend this game to my colleagues.	Eu recomendaria este jogo para meus colegas.
	Social Interaction (Fu et al., 2009; Savi et al., 2011)	I was able to interact with other players during the game.	I was able to interact with other people during the game.	Eu pude interagir com outras pessoas durante o jogo.
		The game promotes cooperation and/or competition among the players.	The game promotes cooperation and/or competition among the players.	O jogo promove momentos de cooperação e/ou competição entre os jogadores.
		I felt good interacting with other players during the game.	I felt good interacting with other players during the game.	Eu me senti bem interagindo com outras pessoas durante o jogo.
	Fun (Poels et al., 2007, Savi et al., 2011)	I had fun with the game.	I had fun with the game.	Eu me diverti com o jogo.
		Something happened during the game (game elements, competition, etc.) which made me smile.	Something happened during the game that made me smile.	Aconteceu alguma situação durante o jogo que me fez sorrir.
	Focused Attention (Keller, 1987; Wiebe et al., 2014; Savi et al., 2011)	There was something interesting at the beginning of the game that captured my attention.	There was something interesting at the beginning of the game that captured my attention.	Houve algo interessante no início do jogo que capturou minha atenção.
		I was so involved in my gaming task that I lost track of time.	I was so involved in the game that I lost track of time.	Eu estava tão envolvido no jogo que eu perdi a noção do tempo.
		I forgot about my immediate surroundings while playing this game.	--	--
	Relevance (Keller, 1987; Savi et al., 2011)	The game contents are relevant to my interests.	The game's content is of my interest.	O conteúdo do jogo me interessa.
		It is clear to me how the contents of the game are related to the course.	It is clear to me how the contents of the game are related to the course.	É claro para mim como o conteúdo do jogo está relacionado com a disciplina.
		This game is an adequate teaching method for this course.	I learned content of the course with this game.	Eu aprendi conteúdo da disciplina com este jogo.
		I prefer learning with this game to learning through other ways (e.g. other teaching methods).	I prefer learning with this game than through other ways (e.g. expositive lectures given by the teacher).	Eu prefiro aprender com este jogo do que de outra forma (p.ex. aula no quadro pelo professor).
	Perceived Learning (Sindre & Moody, 2003; Savi et al., 2011)	The game contributed to my learning in this course.	Descriptive question "What did you learn playing the game?"	Pergunta descritiva" O que você aprendeu jogando esse jogo?"
		The game allowed for efficient learning compared with other activities in the course.	--	--

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).

Items measuring the achievement of 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. Differently to the MEEGA+ measurement instrument in which the achievement of the learning goals is measured through a self-assessment of the participants, this measurement is done as part of the MEEGA+KIDS questionnaire through a set of multiple-choice questions assessing the achievement of the learning goals. These multiple choice questions have to be carefully defined in accordance to the specific learning goals and the competence level to be achieved in accordance with the revised version of Bloom's taxonomy (Anderson, Krathwohl, & Bloom, 2001)(Simpson, 1972)(Krathwohl, Bloom, & Masia, 1973). As part of a self-assessment of the perception of the learning effect of the game we included also a descriptive question "What did you learn playing the game?"

Templates of the MEEGA+KIDS questionnaire are available in English and Brazilian Portuguese on the site: <http://www.gqs.ufsc.br/meega-a-model-for-evaluating-educational-games/>

An example instantiation of the template is presented in Appendix A (in Brazilian Portuguese).

Analysis of the data collected. Following the MEEGA+KIDS 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 in secondary school. 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>* have a good usability?

Player Experience

AQ2: Does the *<name of the evaluated game>* provide 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+KIDS 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 positive player experience?	Measures of central tendency (median, average and frequency of responses); Graphical visualization (frequency charts)	Data collected through the MEEGA+ measurement instrument on the player 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.

4. Conclusions

In this technical report, we propose the a preliminary version of the MEEGA+KIDS model, a customization of MEEGA+, a quality evaluation model of educational games used as instructional strategy for computing education. The model focuses on the quality evaluation of educational games (including digital as well as non-digital games such as card or board games) in terms of player experience and perceived learning in the context of computing education in secondary school. As next steps of our research we are planning a study in order to evaluate of the model in terms of reliability and validity.

Acknowledgments

We would like to thank all the students that participate in the participatory design.

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

References

- 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., & Hailey, T. (2011). The role of psychology in understanding the impact of computer games. *Entertainment Computing*, 2, 69–74.
- Boyle, E. A., Hailey, 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.
- Çiftçi, 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 Programa), Universidade de Caxias do Sul.
- 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.
- 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.

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., & Ijsselstein, 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. In: *Acm 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). Força dos Sinônimos: Um Jogo da Força Multiplayer para o ensino de Tesouros 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.
- Spinuzzi, C. The methodology of participatory design. *Technical Communication*, 52(2), 2005.
- 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*. 2 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.

Appendix A. Example questionnaire for the game SplashCode

MEEGA+KIDS

Questionário para a avaliação da qualidade de jogos não-digitais na educação básica

Nome do jogo: SplashCode

Gostaríamos que você respondesse as questões abaixo sobre a sua percepção da qualidade do jogo para nos ajudar a melhorá-lo. Todos os dados são coletados anonimamente e somente serão utilizados no contexto desta pesquisa. Algumas fotografias poderão ser feitas como registro desta atividade, mas não serão publicadas em nenhum local sem autorização.

Nome do pesquisador/professor: Christiane G. von Wangenheim

Local e data: Florianópolis 20/08/2018

Informações Demográficas	
Instituição	Escola xxx
Ano/Turma ou Curso (no caso de escola técnica)	8 e 9 ano
Disciplina	Interdisciplinar
Idade	
Gênero	<input type="checkbox"/> Menino <input type="checkbox"/> Menina
Com que frequência você joga videogames?	<input type="checkbox"/> Nunca <input type="checkbox"/> Raramente <input type="checkbox"/> Pelo menos uma vez por mês <input type="checkbox"/> Pelo menos uma vez por semana <input type="checkbox"/> Todos os dias
Com que frequência você joga jogos de cartas, tabuleiro, etc.?	<input type="checkbox"/> Nunca <input type="checkbox"/> Raramente <input type="checkbox"/> Pelo menos uma vez por mês <input type="checkbox"/> Pelo menos uma vez por semana <input type="checkbox"/> Todos os dias

Por favor, **marque uma opção** da sua opinião sobre cada afirmação abaixo.

Afirmções	Discordo totalmente	Discordo	Nem discordo, nem concordo	Concordo	Concordo totalmente
O design do jogo é atraente (tabuleiro, cartas, etc.).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Os textos, cores e fontes do material do jogo combinam.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
O tamanho e estilo de letras utilizadas no jogo são legíveis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
As cores utilizadas no jogo são compreensíveis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Aprender a jogar este jogo foi fácil para mim.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
As regras do jogo são claras e compreensíveis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu considero que o jogo é fácil de jogar.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A organização do conteúdo me ajudou a estar confiante de que eu iria aprender com este jogo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Este jogo é desafiador suficiente para mim.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
O jogo oferece novos desafios (novos obstáculos, situações ou variações) com um ritmo adequado.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
O jogo não se torna monótono nas suas tarefas (repetitivo ou com tarefas chatas).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Completar as tarefas do jogo me deu um sentimento de satisfação.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
É devido ao meu esforço pessoal que eu consigo avançar no jogo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Me sinto satisfeito com as coisas que aprendi no jogo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu recomendaria este jogo para meus colegas.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu pude interagir com outras pessoas durante o jogo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
O jogo promove momentos de cooperação e/ou competição entre os jogadores.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu me senti bem interagindo com outras pessoas durante o jogo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu me diverti com o jogo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Aconteceu alguma situação durante o jogo que me fez sorrir.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Houve algo interessante no início do jogo que capturou minha atenção.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu estava tão envolvido no jogo que eu perdi a noção do tempo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
O conteúdo do jogo me interessa.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
É claro para mim como o conteúdo do jogo está relacionado com a disciplina.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu aprendi conteúdo da disciplina com este jogo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu prefiro aprender com este jogo do que de outra forma (p.ex. aula no quadro pelo professor).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Marque a resposta correta. Cada questão tem uma ÚNICA resposta correta.

1. Um algoritmo:

- ☐ é a parte interna de um computador
- ☐ são os passos necessários e ordenados para realizar uma tarefa
- ☐ são as imagens que aparecem na tela do monitor
- ☐ é um jogo de computador

2. Um programa de computador:

- ☐ são os dispositivos (como teclado, mouse, etc.) que você usa para interagir com o computador
- ☐ são as partes internas do computador (como processador, memória, disco rígido, etc.)
- ☐ é a tradução de um algoritmo em instruções que o computador entende
- ☐ é o conjunto dos comandos para jogar um jogo no computador

3. Considere os seguintes passos de um algoritmo para escovar os dentes:

1. Enxaguar com água.
2. Limpar a escova de dentes.
3. Pegar a escova de dentes.
4. Escovar os dentes.

5. Guardar a escova de dentes.
6. Colocar a pasta de dentes na escova.

Qual a ordem correta dos passos do algoritmo para escovar os dentes?
(Responda da seguinte forma: por exemplo: 1-2-3-4-5-6).

Resposta: __ - __ - __ - __ - __ - __

O que você aprendeu jogando esse jogo?	
O que você gostou no jogo?	
O que você achou ruim?	
Mais algum comentário?	

Muito obrigado pela sua contribuição!