

Research Article

Evaluating Multiple Aspects of Educational Computer Games: Literature Review and Case Study

**Juvane Nunes Marciano, Leonardo Cunha de Miranda,
and Erica Esteves Cunha de Miranda**

*Department of Informatics and Applied Mathematics, Federal University of Rio Grande do Norte (UFRN),
59078-970 Natal, RN, Brazil*

Correspondence should be addressed to Juvane Nunes Marciano; ju_n7@yahoo.com.br

Received 31 August 2014; Accepted 10 November 2014; Published 7 December 2014

Academic Editor: Hanqiu Sun

Copyright © 2014 Juvane Nunes Marciano et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Evaluation is an important step in the life cycle of software, once through this practice it is possible to find issues that could compromise the user experience. With educational computer games, the same rule is applied. The use of educational games is increasing, and it is important to verify these tools to provide users with the most adequate learning environments. This verification can be made through the evaluation of multiple aspects of these tools. This work presents a literature review about evaluation of multiple aspects of software, followed by a more specific review focused on multiple aspects of educational computer games. Then, a case study is presented, in which an evaluation method is applied with an educational computer game, aiming to verify the positives and the issues to be improved in the game. The reviews and the description of the process to use the method intend to help and guide other researchers to choose evaluation methods that can fit their own context and needs.

1. Introduction

Interface evaluation of computer systems in general includes usability as a main aspect, and when it comes to the field of educational computer games, other aspects arise, like playability, mechanics, story, and educational content. To accomplish these evaluations, manual, semiautomatic, and automatic methods can be used, and each one of them can detect different interface issues. The interface evaluation brings for the developers several questions and points of view about a system, which they can use to improve it and solve problems that improve the experience of the end user. The accomplishment of an interface evaluation is fundamental to analyze the user experience aspects of a given system.

The use of technology to enhance learning experiences grows increasingly in the field of language learning; these educational applications are called Computer Assisted Language Learning (CALL) tools or applications. CALL applications can be tutor systems, questionnaires, multiuser sharing experience, and also games. A CALL game is a game that can support language students to learn an idiom, providing a fun

environment but still presenting educational content about the focus of the learning. These games are even more popular nowadays, for that it is important to find ways to evaluate them, aiming to find out their positives and the issues to be treated, to make them more enjoyable and useful as a learning tool for the students. Our main objective with this work is to find a method that evaluates satisfactorily a CALL game.

This work presents a literature review about evaluation of multiple aspects of software, reporting evaluation methods and applications. The review has two main steps, a more general overview of methods and then a search for more specific methods to our context. The study aims to comprehend the context of use of the different evaluation techniques—general and specific—with the goal to be able to choose and even adapt a method to be applied in our context, which is language learning supported by computer games. We intend that this adaptation process can be achieved by others through the steps detailed in the present work, and then other researchers can obtain methods that fit more their needs and raise interesting and pertinent questions to their research context. Then, we present the usage of a method to evaluate

multiple interface aspects of educational computer games, with the accomplishment of a case study with an educational game.

The work is organized as follows: Section 2 presents the literature review; Section 3 presents the case study; Section 4 presents the lessons learned; and Section 5 concludes the paper.

2. Literature Review

The literature review is a fundamental part of this work, because it is from this that the analysis of methods is made, with applications and results. This review is divided into two parts. First, we searched for surveys that present several kinds of interface evaluation and overviews with a more general scope of the theme. Then, a search with a more specific scope was made, aiming to find a method that suits better our needs, which was to evaluate educational game.

2.1. Interface Evaluation in General. The search methodology includes searching works in main online scientific libraries and bibliographic databases, using certain keywords and concatenations of them. The used libraries are ACM DL, IEEE Xplore, Springer, Elsevier, Scielo, Scopus, ISI Web of Knowledge, and also Google Scholar. The keywords searched are a permutation of the following words (presented in this paper in alphabetical order): “evaluation”, “interface”, “literature review”, “state of the art”, “survey”, and “usability”. After the search was accomplished, works were selected to compose this review, based first on the abstracts and, second, on a nimble reading of them. The selected works were then read carefully, analyzed, and classified according to methodology, aspects, and requirement of users in the process.

Ivory and Hearst [1] present the state of the art about automating methods of usability evaluation of user interfaces, exhibiting a taxonomy that highlights the role of automation in this context and to compare several methods. 132 methods are evaluated, applied between WIMP and web interfaces. They highlight that user evaluation is fundamental but can be increased and complemented by automated methods and also contribute to lower costs and time of the evaluation as a whole. Følstad et al. [2] present a survey on practical usability evaluation, using a methodology that includes observation and qualitative interviews. The survey respondents were 224 usability practitioners with experience, who answered questions about their experiences in usability evaluations they have participated in. The conclusion is that the evaluation techniques are often adapted by the practitioners, for the given context.

Hilbert and Redmiles [3] performed a survey about user interface events, focused on computer-aided techniques, in which they present a framework to help on approaches’ categorization and comparison. The authors concluded that computer-aided techniques for evaluation of interface can bring out the overall problems that can be detected by an automatic method but lack points that can only be perceived by the analysis of a human being, so the techniques are complementary. Grossman et al. [4] perform a survey about

learnability, which is an important aspect of usability, that represents how easy it is to the users to learn how to use a system. A traditional think-aloud method is compared to a question-suggestion method, using 10 undergraduate architecture students, volunteers, to evaluate AutoCAD’s learnability. The results made the authors develop a new evaluation protocol to identify learnability issues, and they conclude that the methodologies can be used together.

When this first search was accomplished, we perceived the need to search for more specific methods to fit our context, which is to evaluate an educational game. In the next subsection, we present a literature review focused on methods for evaluation of educational games.

2.2. Evaluation of Multiple Aspects of Educational Games. This overview presented next allows us to get knowledge about evaluation in the context of educational games, taking into account several aspects of interface to be evaluated. Once this is a very specific field, there is not a large amount of work related; still, this does not intend to be an exhaustive list of related works, but a selection of the main research and techniques currently used in this context. Different from the works cited in the introduction, which are surveys about interface evaluation in general, this review is focused on evaluating multiple/different interface aspects of educational tools. The search methodology followed was the same as above, but with different keywords, which were “aspects”, “case study”, “educational game”, “evaluation”, “interface”, and “playability”. The criteria of choice were similar to the first search too.

Omar and Jaafar [5] aimed to reach a heuristic method to evaluate educational games, with focus on Malaysian ones. With their method, they intended to need few persons to evaluate, including designers, developers and usability professionals. They use heuristics based on Nielsen’s, relating them to the context of educational games, for evaluation of playability. Another work by the same authors [6] brings a set of heuristics and a framework of five steps to apply them in an interface evaluation using prospective and end users, with heuristics focused on educational games that were divided into five aspects of educational games, that is, interface, educational/pedagogical, content, multimedia, and playability. In a later work of these authors [7], Omar et al. improved the previous work presenting a tool—AHP_HeGES—to facilitate this kind of evaluation during the game’s development process. The methods are not validated in those works.

Desurvire et al. [8] bring a set of heuristics to evaluate playability of video, computer, and board games, based on other heuristics from the literature. The method, called HEP, is to be applied by an expert evaluator. Their method is applied in the work with the development of a game which is used to evaluate the method. They also made a user centered evaluation, using four prospective users, in order to compare its results with the HEP method and concluded that HEP found a considerable higher number of issues over the user centric approach, although the latter found more specific issues for the game.

Liao and Shen [9] present the interface evaluation of a computer game-based learning material, called *Crazy Machines 2*, using heuristics and user centric approaches to determine if the game was developed under usability principles. The heuristics used are present in a work referred to by the authors; and the user centric evaluation was made with five research students from courses as education and communication technologies. Korhonen et al. [10] accomplish a comparison work between two playability heuristic sets. The evaluation was performed with eight persons from game industry and academy, divided in four pairs. The results of the study point out that the use of heuristics guided the evaluators well, including during the development phase; a lower point of heuristics, according to the evaluators, is that since there are a lot of kinds of games, it is difficult to make heuristics that work for all of them; on the other side, too much specific heuristics may restrict their use. Another interesting finding is that it is hard for the evaluators to have a real game experience during the evaluations, once they have to be alert to the playability problems that may occur. The authors concluded that this is a valid method but needs improvement for the games area.

Pinelle et al. [11] develop a new set of heuristics to evaluate usability of game interfaces. They base their work on Desurvire's [8] and claim that the HEP heuristics have limited coverage of usability issues; then, they presented a set of heuristics focused on usability, made by the authors. To do that, the authors analyzed a list of 108 computer game reviews, and once the problems were identified and classified, heuristics would be developed, that would be the opposite of the given problem. They reached a number of 12 categories of problems and developed a list of ten heuristics, highlighting that they are based on Nielsen's. The evaluators were five videogame players with experience in usability evaluation.

Barcelos et al. [12] propose another new set of heuristics for evaluation of digital games and validate them comparing to another existing set. Their set is a compilation of some others, including Pinelle et al. [11], and what they aimed to do was to create a shorter and more direct set of heuristics that could be as efficient as the larger ones that they were based. For the evaluation, the games Earth 2160 and Outlive were chosen, both from the category Real Time Strategy of games. The evaluators were 34 HCI students of an upper course of digital games; they were divided in groups in which some of them received the list of 18 heuristics developed by the authors and the others received the 40 Federoff's [13] heuristics. They conclude, by the results of the evaluation, that the performance of both sets is very similar.

Grace et al. [14] provide an initial evaluation of a language learning game for Mandarin. The evaluation was made with 21 potential users of the game, called Polyglot Cubed. By means of questionnaires, the authors analyzed the response of the users and were able to detect some interface issues with the game, as problems with selecting objects while playing and seeing pictographic representations of words. Serrano et al. [15] accomplish an evaluation of an educational game through observation of users' interaction. While the players do play the game, data is collected automatically and analyzed to extract relevant issues. They proposed a framework and

TABLE 1: Surveys about multiple interface aspects evaluation methods.

Work	Evaluation	
	Based on	Focus
[1]	Automating methods	Usability of user interfaces
[2]	Observation and qualitative interviews	Usability
[3]	Computer-aided techniques	User interface events
[4]	Think-aloud method and question-suggestion method	Learnability

performed a case study with a basic math game to validate it.

Lazareck et al. [16] performed an evaluation of an educational game for 13–15-year-olds, which teaches about biotics and hygiene. The evaluation was made with 129 pupils of three different UK schools, in a user centric approach that used a questionnaire the users answered after playing the game. The questions addressed mainly to usability and playability, and the answers were given in varied ways, since yes/no until comments they wanted to make about the game, and then the responses were analyzed. Hersh and Leporini [17] performed a study about accessibility in educational games for disabled students. The study was applied on disabled students and people that are close to them, as family and educators, by means of a questionnaire. The results of the study gave the authors answers for them to develop guidelines and recommendations about the development of this kind of game for these target users. The guidelines, beyond accessibility, also evaluate pedagogical aspects and playability/usability.

Park and Kim [18] present a study about accessibility for disabled people, in serious games for mobile environment. The authors suggest guidelines for development of this kind of game taking into account the following aspects: definition of game goals; planning game scenario; planning game components; planning rewards; planning training progress; and planning of training effects and achievement. Each one of these aspects has a guideline associated. Torrente et al. [19] bring a study about accessibility of educational games for blind students. They developed three different eyes-free interfaces, applied on a point-and-click game. The cyclical navigation system provides a two-level interaction using right and left arrows and buttons “enter” and “escape”, in which the elements are selected and the options are passed through sounds for the player. The second, called sonar, uses the mouse and sounds to indicate that the cursor is near an object. And the third is a natural language commands interface, with typed commands. An evaluation of the three interfaces was made analyzing them according to usability, engagement, and additional cost. They also mention a pilot study with ten blind students, though there are no described results about this study.

2.3. Synthesis and Discussion. In this subsection, we summarize the works of the literature review in Tables 1 and 2.

TABLE 2: Summary of the works of the literature review with focus on educational games.

Work	Based on	Focus	Evaluation	Applied	End users
[5]	Heuristics	Playability		No	No
[6]	Heuristics	Interface, pedagogical, content, playability, and multimedia		No	Yes
[7]	Heuristics and user centric	Interface, pedagogical, content, playability, and multimedia		No	Yes
[8]	Heuristics and user centric	Playability, story, mechanics, and usability		Yes	No
[9]	Heuristics and user centric	Interface, rules, story, and social		Yes	Yes
[10]	Heuristics	Playability		Yes	No
[11]	Heuristics	Usability		Yes	Yes
[12]	Heuristics	Playability		Yes	No
[14]	User centric	Usability		Yes	Yes
[15]	Analysis	Usability and pedagogical		Yes	Yes
[16]	User centric	Usability and playability		Yes	Yes
[17]	Guidelines	Usability, accessibility, playability, and pedagogical		Yes	Yes
[18]	Guidelines	Accessibility		No	No
[19]	Analysis	Accessibility		Yes	Yes

Table 1 summarizes the works of literature previously presented about multiple interface aspects evaluation. This table presents the column names: “Based on” which indicates what methods are cited in the work and “Focus” that indicates the main focus of the evaluation methods described. Table 2 summarizes the works of literature previously presented, with focus on evaluation of educational games. This table presents the column names: “Based on” which indicates if the method is based on heuristics, user centric, analysis, and/or guidelines; “Focus” that indicates the evaluation focus, that is, usability, accessibility, playability, interface, pedagogical, social, multimedia, content, story, rules, and/or mechanics; “Applied” and “End Users” that indicate, respectively, if the method was applied and if the method needs end users during the evaluation process, that is, yes or no.

In Table 2, it is possible to notice a trend of heuristic evaluation followed by user centric evaluations. The heuristic evaluation has the advantage of being flexible and can also be applied with prospective and end users. Related to the focus, the most evaluated are playability, usability, and pedagogical aspects. The works that evaluate accessibility are specific and use techniques as guidelines and comparative analysis. The majority of works apply their techniques to validate them practically. Some of the techniques presented are just exhibited and explained but not applied in the same work.

Until this point, we were able to know several methods to evaluate multiple aspects of interface, through the review of surveys which present several methods to evaluate interface aspects of a given tool. Then, concluding that a more general review was important, but to evaluate precisely the aspects of an educational game, it was needed to search for more specific methods, and then the second review was accomplished. After the accomplishment of these literature reviews, we could reach some conclusions about the methods that could be used to evaluate an educational game with the most satisfactory results for our context.

The second review brought some interesting works which could be used to evaluate an educational game. After

analyzing these methods, we chose to use the PHEG methodology by Omar and Jaafar [6], which is very interesting because it evaluates, in fact, multiple interface aspects concerning specifically educational games. This method is heuristics based and evaluates, from an interaction point of view, five aspects of educational games: interface, educational/pedagogical, content, multimedia, and playability. Their method is explained in a little more detail in the next section.

3. Case Study

In this Section, we present an overview of the educational game which will be evaluated in the case study, that is, Karuchā Ships Invaders game, followed by the methodology applied for the evaluation, and the results obtained.

3.1. Karuchā Ships Invaders CALL Game. The Japanese language is a very challenging idiom to learn, as it has three main alphabets that are all different from the Roman alphabet used in most languages around the world. There are kanji, ideograms that were originally Chinese but were imported and adapted to Japanese language; they represent ideas and one kanji alone can have several pronunciations. The other two Japanese alphabets are hiragana and katakana (which together are called kana); they are different from kanji since they are syllabic alphabets; the former is the most basic, used for writing words that do not have a kanji, also for adjectives and verb endings, and as a subtitle for rare or unknown kanji (the subtitle is called furigana). Hiragana is the first taught alphabet to anyone learning Japanese language, in or outside Japan. Katakana is an alphabet similar to hiragana in its basis, but the symbols are all different, and it is used to adapt foreign words to the Japanese idiom, and in some cases for highlighting a certain word or onomatopoeia. In Japan, they also use the Roman alphabet—which is used in most languages—but barely, for example, in acronyms.

In this context, Karuchā Ships Invaders is a CALL game (available for free download at <http://www.karucha.pairg.dimap.ufrn.br/>) developed by our research team and presented in detail in [20], with the goal of supporting Japanese language learning, by means of presenting the most basic Japanese alphabet, that is, hiragana, plus some cultural words related to Japanese culture. The concept of the game is a story where Japanese ships are heading to Brazil—represented by a city with Brazilian elements on the game screen—and need help to land; this help has to come from Brazilians that know the basics of the Japanese language, to pass the commands to the ships.

The gameplay is space invaders style, where the ships act like “invaders.” Each ship brings a Japanese letter on it or a picture representing a cultural Japanese word written with the basic alphabet; and the player has to type, on the keyboard, the corresponding pronunciation using the Roman alphabet and then press the Enter key or Backspace key. If the typing is correct, a laser will be shot and embrace the ship, helping it to land safely; however, if a ship is not guided—the player does not type its pronounce—it will fall down and cause damage to the city below. Figure 1 shows six screenshots of the gameplay: level choice screen (Figure 1(a)), gameplay screens (Figures 1(b) and 1(c)), context-help menu (Figure 1(d)), educational stories (Figure 1(e)), and the main menu screen (Figure 1(f)).

To present/teach new content to the player, the game exhibits stories protagonized by two Brazilians that went to Japan to learn about its culture; the stories are reports to their friends in Brazil, to help them be able to guide the ships that come to land. The game affords three difficulty modes, that is, easy, normal, and hard, each one increasing velocity to the ships’ fall. Still, each difficulty mode contains 30 levels to be passed, with increasing number of hiragana characters and cultural words to be learned. Hiragana is presented in alphabetical order and is exhibited in all levels but 15 and 30, which are levels that only contain cultural words called the bosses of the game. Also, every three levels, the third combines hiragana characters with cultural words or bosses.

The Karuchā Ships Invaders game was developed to provide an immersive environment for the users—with a focus on, but not limited to, Brazilian students of the Japanese language—presenting the Brazilian city that receives Japanese visitors. In Brazil side, there are the city elements that remind them of Brazil, like, for example, the statue of Christ the Redeemer, Maracanã Stadium, favela, and the famous Copacabana boardwalk on the beach. Japan is represented on the ships, which bring a Japanese flag in their front, plus the letters themselves that are unique of the Japanese writing system. Multimedia resources also contribute to the immersive nature of the game, using, for example, sounds and images that remind them of both Japan and Brazil.

3.2. Methodology and Application of the Method. The methodology chosen to be applied in this study is, initially, heuristic based, with the developers being evaluators, for detecting overall interface issues, based on the work of Omar and Jaafar [6]. The choice of this work in particular was made because it has a set of heuristics focused on evaluation

of educational computer games, which is our focus too. Also, these heuristics are divided into five axes related to the user interaction experience with focus on the educational purposes of a given game, allowing for evaluating multiple aspects of the tool with one method. Still, this work [6] is a result of several researches [5, 7, 21–25] focused on this subject-matter. The work [6] brings the set of heuristics and also provides a framework with five steps to apply, represented in Figure 2. From now on, we describe the steps of the method, with focus in detail on our process of application focusing on the theme of language learning, though aiming to guide other researchers who want to apply it even for other educational focuses.

The first step consists in developing a questionnaire based on the heuristics that suits the context of the game to be evaluated. After that, an identification of possible evaluators is made, between experts and prospective and end users. The third step plants the way the evaluation will be conducted, as location, time, and form of presentation of questionnaire. The fourth step designs the evaluation as a whole, determining which will be made in the available time and the order to be followed. The fifth step is related to the analysis after the evaluation and dictates what will be made with the results. The steps conducted in this study are presented in this section.

The heuristics used in the expert evaluation—Playability Heuristics for Educational Games (PHEG)—are divided in five main aspects of an educational game, that is, (i) interface, which includes factors as interactivity, navigation, design, and consistency (UI has ten heuristics); (ii) educational/pedagogical, which evaluates goal and objective, challenge, feedback, and player control (ED has ten heuristics); (iii) content, which contains aspects related to the educational material provided (CO has eight heuristics); (iv) multimedia, which addresses the use of text, audio, animations, and all kinds of media, if they are appropriate (MM has eight heuristics); and (v) playability, which evaluates questions such as balance, pace, control of the player over the game, and adequacy of levels (PL has seven heuristics), in a total number of 43 heuristics.

The PHEG evaluation can be performed by experts or by prospective/end users. For the evaluation, we developed a set of questions based on the basis heuristics of five aspects previously described, for guiding the evaluators in the process. The method requires two or more questions per aspect, but it is not mandatory to develop questions over all the heuristics. The questions of this study were developed based on the heuristics that fit the context and needs of Karuchā Ships Invaders CALL game. Table 3 presents the heuristics used as basis for the development of the questions (cf. [6]); it is important to highlight that in this table we do not present all 43 heuristics, only the ones which we used to develop questions for our work.

Table 4 presents the questions developed specially for the evaluation of Karuchā Ships Invaders CALL game, which represents the first step of the framework. These questions were developed by our team of researchers, to be applied with the evaluators, in this case, interface/interaction experts.



FIGURE 1: Different moments of the Karuchā Ships Invaders educational game: (a) level choice, (b) falling cultural picture, (c) ship being helped to land, (d) example of context-sensitive help (level 24), (e) educational story, and (f) menu screen.

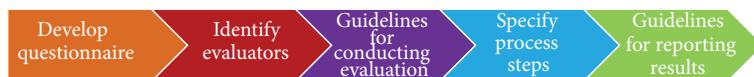


FIGURE 2: Five steps of the evaluation method.

3.3. Results. The evaluation was made using HCI experts as respondents for the questions developed about Karuchā Ships Invaders CALL game. The answers were given subjectively, since the method used is an adaptation for Omar and Jaafar [6]. It was chosen to analyze the data beyond objective

numbers as answer. Table 5 presents the answers for the 28 heuristic based questions previously defined (Table 4).

The results provided, that is, the answers by the experts (Table 5), are the data to be analyzed regarding Karuchā Ships Invaders' interface. In the fourth step of the process,

TABLE 3: Basis heuristics for the development of the questions (cf. [6]).

Aspect	Code	Brief description	Heuristic
Interface	UI1	Uses aesthetic and minimalist design	
	UI3	The uses of space, color, and text are according to the principles of screen design	
	UI4	The uses of text, color, and font follow the principles of readability	
	UI6	The interactivity of the game is suitable to learners level	
Educational/pedagogical	UI9	Provide specific and self-identified key for specific task (e.g., exit, glossary, main, and objective)	
	ED1	Clear goal and learning objectives	
	ED2	The activities are interesting and engaging	
	ED4	Can be used as self-directed learning tools	
	ED9	Offers the ability to select the level of difficulty in games	
Content	ED10	Ability to work in their own pace	
	CO1	Reliable content with correct flow	
	CO3	Navigation of content is easy and accurate	
	CO4	Supporting materials are sufficient and relevant	
	CO5	Materials are interesting and engaging	
Multimedia	CO6	Players able to understand the learning goal	
	MM2	Combination of multimedia elements is adequate	
	MM3	The presentation of multimedia elements is well managed	
	MM4	Suitability of multimedia elements for specific use	
	MM6	The use of multimedia elements supports meaningfully the text provided	
Playability	MM7	The quality of multimedia elements (e.g., text, image, animation, video, and sound) used is acceptable	
	PL3	The pace of the game is in balance	
	PL4	Players are able to control the game	
	PL5	Progress of the game can be seen at any time	
	PL7	Challenge is adequate, that is, not too easy and not too difficult	

it was decided to apply the questionnaire, then analyze the responses, and then generate quantitative results based on these responses, leading to detection of specific number of problems and the corresponding severity rating. Table 6 presents the severity rating for each one of the five aspects evaluated; besides the severity ratings, there is an indication of what question is detected the issue. This table is a representation of what was decided to be made in the fifth step, based on the framework.

It is interesting to highlight that no problem with severity rating 3 or 4 was found in this evaluation. This may have to do with the fact that Karuchā Ships Invaders was developed over considerable previous research [20], providing a better sight of what was desirable to be developed.

4. Lessons Learned

This work brings a noteworthy contribution to the field of evaluation of game interfaces in general, as it brings a literature review that can guide researchers that aim to accomplish this kind of evaluation. Still, the evaluation that was performed for Karuchā Ships Invaders also shows that methods can be adapted for the needs and context of each game, focusing on educational computer games. The framework and heuristics used are adaptive and can be applied

for experts or for prospective/end users, depending on the questions developed based on these heuristics.

One important thing learned in this process was about the recommendation of the method, to apply with experts or end users. We chose to use experts only and perceived that was not enough. In our case, the evaluators (who answered the questionnaire) were domain experts (in Japanese language learning) and HCI experts, but we did not ask for an end user to participate in the evaluation (a Japanese student with no prerelation to the game). Though, this was our objective, as this evaluation was performed to detect issues that could embarrass the user experience, and to reach this goal, the experts' evaluation was adequate. Nevertheless, this choice lowered the diversification of answers we could obtain, and this was perceived in a further evaluation with focus on motivational questions of the same game, using end users [26, 27].

In the work [27] we accomplished an evaluation of the motivational aspects of the game Karuchā Ships Invaders, with end users as the evaluators, answering a questionnaire after having the experience of playing the game. In this evaluation, several issues related to the questions presented here were raised by the users, issues that were not thought of or perceived by the domain experts or by the HCI experts, once they did not pursuit the point of view of a real student who would use the game to learn Japanese. Though

TABLE 4: Questions defined to assess multiple interface aspects of the Karuchā Ships Invaders CALL game.

Aspect	Basis heuristic (Table 2)	Question
Interface	UI1	(Q1) Graphical elements are consistent and adequate?
	UI3	(Q2) Screen space is well distributed (each element has enough space)?
	UI4	(Q3) Font size for romaji is adequate?
	UI6	(Q4) Font size for Japanese characters is adequate?
	UI9	(Q5) Interaction mode is adequate to the audience?
		(Q6) Access buttons for different functions and features are easily identified?
Education/pedagogical	ED1	(Q7) Menu division is adequate?
	ED2	(Q8) Learning goals are clear?
	ED4	(Q9) Gameplay is incentive and engaging?
		(Q10) Is it possible to obtain new considerable knowledge about Japanese language by just playing the game?
	ED9	(Q11) Does the game provide choice of difficulty level?
	ED10	(Q12) The difficulty levels are adequate, that is, too easy/hard, or discrepant?
Content	CO1	(Q13) The game suits to different learning paces of the students?
	CO3	(Q14) The content is presented in gradual and coherent way?
	CO4	(Q15) Is it easy to access support content, that is, stories?
	CO5	(Q16) The educative stories bring the content in a comprehensive way?
	CO6	(Q17) The educative stories do provide enough background for the player to win the game?
		(Q18) Is it interesting/engaging to learn through the stories?
Multimedia	MM2	(Q19) Is it clear for the player the content to be learned along the game?
	MM3	(Q20) The multimedia elements (e.g. sounds, images, and text) are in the context of the game?
		(Q21) The presentation of multimedia content follows an adequate flow? (Is there too much or a lack of multimedia)
		(Q22) The use of the following multimedia elements is adequate?
	MM4	Images in the stories
		Text in the stories
Playability		Ambient music
		Sounds (e.g. laser, shot, and explosion)
		Images in level transitions
	MM6	(Q23) The multimedia elements are significant for the gameplay and learning (or are dispensable)?
	MM7	(Q24) The quality of multimedia elements is acceptable?
	PL3	(Q25) The game flow is adequate?
Playability	PL4	(Q26) The controls of the game are adequate (or are tiring, hard, non-intuitive)?
	PL5	(Q27) Is it easy for the player to perceive or access the current status in the game at any moment?
	PL7	(Q28) The challenge level is adequate to the audience, that is, not too easy and not too difficult?

the evaluation presented in the present work has its value, and with these results we could present a better game to the users in the motivational evaluation. Even though they pointed new issues, the issues treated in the present work were not cited by the users. This shows that the present evaluation was important to provide a better experience for the users during their own evaluation, without concerns about basic interface issues, which were previously detected by the present work.

For the game Karuchā Ships Invaders, there are some interesting issues pointed by the evaluation. The answers for the questionnaire also raise discussions, which follow, about each one of the five analyzed aspects of usability. The interface analysis shows that there are some issues such as lack of personalization for features as font size and lack of buttons

and configuration elements at the time a level is being played; this analysis also raises questions about accessibility of the game, once there are no customizable options to support players with special needs.

For the educational/pedagogical aspect, the issues are about the difficulty modes, which do not increase content to be learned in each mode. This question can be answered analyzing the sort of the game, which is "invaders"; this kind of game has the characteristic of increasing speed as the difficulty increases, and related to new content, it is represented by the 30 levels, where each one of them brings new characters or words for the player to learn. Besides, there is a discussion about the learning goal being clear to the player, since the content is not provided at once in

TABLE 5: Answers for the heuristic based questions.

Aspect	Answer
Interface	(Q1) The graphical elements are consistent together, and also are in the context of the learning proposal of the game.
	(Q2) The game screen could be increased in height, for the elements to have more space to fall. Though this change could require other adjustments about resolution.
	(Q3) The romaji fonts along the game are in an acceptable size, but the game should provide personalization for this feature, for possible players with sight issues.
	(Q4) The characters inside the ships are in a size sufficient enough to be seen clearly. Though, it is interesting to give the player the opportunity to personalize this feature.
	(Q5) Considering the audience as learners with abilities to see the elements on the screen and type commands on the keyboard, the interaction mode is adequate.
	(Q6) In the menu screen and inside the menus, the elements are identifiable, but when the game is on, there is no button that indicates pause, menu, configurations of sound or screen. These elements should be represented by icons while the game is being played, for the player to have more control over the game.
	(Q7) Yes, the menu options are well divided.
Educational/pedagogical	(Q8) There is not a presentation of the game in the beginning. The help menu provides some information, but it would be interesting to have some kind of initial animation to present the educational and fun goals of the game.
	(Q9) The gameplay is invaders-like, so it can be engaging for those who like this kind of game. Also can be incentive the fact that the player is learning a new language while playing the game.
	(Q10) Considering the player with no knowledge at all about Japanese language, s/he will get considerable knowledge (an entire alphabet and vocabulary—words related to the culture of the country). For the ones who play Karuchā Ships Invaders for training hiragana skills, there are still the cultural words to be learned.
	(Q11) Yes, the game provides three difficulty modes.
	(Q12) The difficulty modes are just different in velocity. Considering that, they are not discrepant. But related to content, there is no increasing based on the difficulty.
	(Q13) The game provides three different difficulty levels. To suit all paces, it is desirable that the game provide some kinds of personalization, which it does not.
	(Q14) Yes. The hiragana content is presented in alphabetical order, and the cultural words are presented just when all the characters that compound it have been presented before. The amount of content by level is also acceptable.
Content	(Q15) The stories become available as the player passes through the levels (one story for each three levels). Once available, the player has the option to read it as many times as needed, in the level choice screen.
	(Q16) Considering the audience is a native English or Portuguese speaker. The content is well divided and exhibited for the player.
	(Q17) The stories present all the content that will be charged in the game, so it provides all the knowledge the player needs to reach the end of Karuchā Ships Invaders.
	(Q18) The stories are displayed in text and images. Maybe animations could be more engaging and less tiring.
	(Q19) The content is only viewed as a whole when the player gets to the end of the game (30 levels).
	(Q20) In general, the multimedia elements follow the context of the game. The music and sounds could be more related to the theme.
	(Q21) The multimedia is in good amount.
Multimedia	(Q22)
	(a) The images in the stories are fundamental to teach the cultural words, once they are the same images that will appear in the game for the player to recognize.
	(b) The text in the stories is adequate and fundamental.
	(c) The ambient music could be more related to Brazil and Japan at the same time.
	(d) The sounds for laser, shot and explosion are common sounds that can be adequate, but they could remind more of the theme (Brazil—Japan)
	(e) The Japanese pilot is adequate to the context.
	(Q23) The use of educational stories is more acceptable for the player when it presents not only plain text, but images and animations; the multimedia elements should not be considered dispensable in this case. Plus, the gameplay becomes more joyful when the game provides an immersive environment related to the theme, what happens in Karuchā Ships Invaders.
(Q24) The elements of multimedia are in good quality.	

TABLE 5: Continued.

Aspect	Answer
Playability	(Q25) The presentation of elements follows a logic order and the number of levels corresponds to the content that is well distributed between each level. The difficulty grows increasingly including the cultural words that can be formed only by the characters that have already been presented.
	(Q26) Considering a desktop environment, the controls feel adequate, once it focuses on the standard keyboard. The controls are intuitive, since the player just has to type the romaji to the symbols and words presented.
	(Q27) The time, difficulty mode and level are clear on the screen; the hit points bar does not indicate a fix number or the amount of loss in a crash. The total points are also not exhibited.
	(Q28) Considering the audience as learners of hiragana that can be from knowing nothing to be conversant to the alphabet; for the former, it can be considered little hard if the player chooses not to read the educative stories that are presented. The cultural words can be a plus challenge for those who only learnt the alphabet itself. The possibility of choice of difficulty mode embraces most of the audience needs.

TABLE 6: Severity rating of problems found.

Aspect	Number	Problems				
		4	3	2	Severity rating	1
Interface	4	—	—	3 (Q3, Q4 & Q6)	1 (Q2)	—
Educational/pedagogical	2	—	—	1 (Q13)	1 (Q8)	—
Content	1	—	—	1 (Q18)	—	—
Multimedia	1	—	—	—	1 (Q20/Q22c/Q22d)	—
Playability	1	—	—	—	1 (Q27)	—

the beginning of the game. The question is if it is really necessary or in an educational and fun point of view it is interesting to the player to discover new content along the game. The purpose of the locked content is to make the game more engaging and interesting.

The analysis of content perceives that the presentation of the stories with only text and pictures may be exhaustive for the player. The suggestion, to make the learning more gamified, is to add animations to the stories; the animations are less tiring to the player and more interesting. After all the game is a new way to learn, with the bid to be different from traditional learning, in which reading is a noted constant and perhaps should be avoided in this approach.

Multimedia aspect brings attention to the sound and music aspect, which are the less related to the context of the entire game. The ambient music reminds only of Japan, whilst the other sounds are way too common; the latter is not a big issue, but fixing it could render the game a more interesting and immersive learning environment. The analysis of playability aspect, as multimedia, was the one that showed the fewest issues of all; the main question was about the possibility of the player to get access to the current statistics of the game at any time, where a lack of this possibility was found. In this case, this lack can be seen as a gamification characteristic, whereas the player is having fun with the game without the preoccupation of looking at statistics when a level is being played; this feature takes this load away from the player, who is concerned about protecting the city and guiding the ships to land, instead of a learning status.

About the technique used to evaluate the interface of Karuchā Ships Invaders CALL game, it was useful to evaluate the five aspects, but there is a lack it does not embrace. The emotional and affective aspects of the interface cannot be evaluated by this set of heuristics. In the context of educational computer games, as seen in this literature review, this aspect is not contemplated in general evaluations.

5. Conclusions

This paper accomplished a literature review divided in two steps, a more general review about evaluation of multiple aspects of user interface and a more specific review about evaluation of educational games. Through this evaluation, it was possible to choose a method that better fits our goal, which was to evaluate an educational game to support language learning focused on the Japanese idiom. The reviews were followed by a study that applied an evaluation technique focused on educational games, using the Karuchā Ships Invaders CALL game as a case study. The performed evaluation embraced the points interface, educational/pedagogical, content, multimedia, and playability and showed that the game interface is in general adequate to its purpose. Still, it can be concluded that the method used is useful but does not embrace certain aspects, such as the player's emotional response to the game.

The main contributions of this work are the literature review, which reports several methods of evaluation that can fit to most general evaluation purposes—the first part of

he literature review—and also presents more methods that can be used to evaluate specifically educational games. The other contributions of this work are the lessons learned through the application of the method, which was good to this purpose but showed that other aspects that deserved to be evaluated are not covered by only one method and also not by only one point of view (the question discussed about not having end users in the process).

As future work, it is intended to do a hedonic evaluation of the game. Also, it is pointed to improve the interface of the Karuchā Ships Invaders CALL game based on the results of this study, plus doing an evaluation that focuses on the learning value of the game with end users.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was partially supported by the Brazilian National Council of Scientific and Technological Development (CNPq Grant no. 163408/2012-2); by the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES); and by the Physical Artifacts of Interaction Research Group (PAIRG) at Federal University of Rio Grande do Norte (UFRN), Brazil. The authors also thank the team of Karuchā Ships Invaders.

References

- [1] M. Y. Ivory and M. A. Hearst, "The state of the art in automating usability evaluation of user interfaces," *ACM Computing Surveys*, vol. 33, no. 4, pp. 470–516, 2001.
- [2] A. Følstad, E. L.-C. Law, and K. Hornbæk, "Analysis in practical usability evaluation: a survey study," in *Proceedings of the 30th ACM Conference on Human Factors in Computing Systems (CHI '12)*, pp. 2127–2136, ACM, May 2012.
- [3] D. M. Hilbert and D. F. Redmiles, "Extracting usability information from user interface events," *ACM Computing Surveys*, vol. 32, no. 4, pp. 384–421, 2000.
- [4] T. Grossman, G. Fitzmaurice, and R. Attar, "A survey of software learnability: metrics, methodologies and guidelines," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*, pp. 649–658, April 2009.
- [5] H. M. Omar and A. Jaafar, "Playability Heuristics Evaluation (PHE) approach for Malaysian educational games," in *Proceedings of the International Symposium on Information Technology (ITSim '08)*, pp. 1–7, Kuala Lumpur, Malaysia, August 2008.
- [6] H. M. Omar and A. Jaafar, "Development and potential analysis of heuristic evaluation for educational computer game (PHEG)," in *Proceedings of the 5th International Conference on Computer Sciences and Convergence Information Technology (ICCIT '10)*, pp. 211–219, Seoul, Republic of Korea, November 2010.
- [7] H. M. Omar, R. Ibrahim, and A. Jaafar, "Methodology to evaluate interface of educational computer game," in *Proceedings of the International Conference on Pattern Analysis and Intelligent Robotics (ICPAIR '11)*, pp. 228–232, June 2011.
- [8] H. Desurvire, M. Caplan, and J. A. Toth, "Using Heuristics to Evaluate the Playability of games," in *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '04)*, pp. 1509–1512, April 2004.
- [9] Y. H. Liao and C.-Y. Shen, "Heuristic evaluation of digital game based learning: a case study," in *Proceedings of the 4th IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning (DIGITEL '12)*, pp. 192–196, Takamatsu, Japan, March 2012.
- [10] H. Korhonen, J. Paavilainen, and H. Saarenpää, "Expert review method in game evaluations: comparison of two playability heuristic sets," in *Proceedings of the 13th International MindTrek (MindTrek '09)*, pp. 74–81, ACM.
- [11] D. Pinelle, N. Wong, and T. Stach, "Heuristic evaluation for games: usability principles for video game design," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*, pp. 1453–1462, April 2008.
- [12] T. S. Barcelos, T. Carvalho, J. Schimiguel, and I. F. Silveira, "Comparative analysis of heuristics for digital games evaluation," in *Proceedings of the IHC+CLIHC 2011*, pp. 187–196, SBC, 2011.
- [13] M. Federoff, *Heuristics and usability guidelines for the creation and evaluation of fun in video games [M.S. thesis]*, Indiana University, 2002.
- [14] L. Grace, M. Castaneda, and J. Ducher, "User testing of a language learning game for Mandarin Chinese," in *Proceedings of the Extended Abstracts on Human Factors in Computing Systems (CHIEA '12)*, pp. 865–868, 2012.
- [15] A. Serrano, E. J. Marchiori, A. del Blanco, J. Torrente, and B. Fernandez-Manjon, "A framework to improve evaluation in educational games," in *Proceedings of the IEEE Global Engineering Education Conference (EDUCON '12)*, pp. 1–8, April 2012.
- [16] L. J. Lazareck, D. Farrell, P. Kostkova, D. M. Lecky, C. A. McNulty, and D. Weerasinghe, "Learning by gaming—evaluation of an online game for children," in *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '10)*, pp. 2951–2954, IEEE, Buenos Aires, Argentina, September 2010.
- [17] M. Hersh and B. Leporini, "Accessibility and usability of educational gaming environments for disabled students," in *Proceedings of the 12th IEEE International Conference on Advanced Learning Technologies (ICALT '12)*, pp. 752–753, July 2012.
- [18] H. J. Park and S. B. Kim, "Guidelines of serious game accessibility for the disabled," in *Proceedings of the International Conference on Information Science and Applications (ICISA '13)*, pp. 1–3, IEEE, 2013.
- [19] J. Torrente, E. J. Marchiori, J. A. Vallejo-Pinto, M. Ortega-Moral, P. Moreno-Ger, and B. Fernandez-Manjon, "Eyes-free interfaces for educational games," in *Proceedings of the International Symposium on Computers in Education (SIIE '12)*, pp. 1–6, IEEE, October 2012.
- [20] J. N. Marciano, A. L. S. Ferreira, A. C. C. Correia, L. C. Miranda, and E. E. C. Miranda, "Invaders: cultural issues on the design/development of a Japanese CALL game made by/to Brazilians," in *XII Simpósio Brasileiro de Jogos e Entretenimento Digital (SBGames '13)*, pp. 172–180, 2013.
- [21] H. Mohamed and A. Jaafar, "Development and potential analysis of heuristic evaluation for educational computer game (PHEG)," in *Proceedings of the 5th International Conference on Computer Sciences and Convergence Information Technology (ICCIT '10)*, pp. 222–227, December 2010.

- [22] H. Omar and A. Jaafar, "Heuristics evaluation in computer games," in *Proceedings of the International Conference on Information Retrieval & Knowledge Management (CAMP '10)*, pp. 188–193, IEEE, Shah Alam, Malaysia, March 2010.
- [23] H. Mohamed-Omar, R. Yusoff, and A. Jaafar, "Quantitive analysis in a heuristic evaluation for usability of educational computer game (UsaECG)," in *Proceedings of the International Conference on Information Retrieval & Knowledge Management (CAMP '12)*, pp. 187–192, Kuala Lumpur, Malaysia, March 2012.
- [24] H. M. Omar and A. Jaafar, "Challenges in the evaluation of educational computer games," in *Proceedings of the International Symposium on Information Technology (ITSim '10)*, pp. 1–6, Kuala Lumpur, Malaysia, June 2010.
- [25] H. M. Omar and A. Jaafar, "AHP-HeGES: tools to evaluate usability of educational computer game (UsaECG)," in *Proceedings of the 2nd International Conference on User Science and Engineering (i-USer '11)*, pp. 73–76, December 2011.
- [26] J. N. Marciano, L. C. Miranda, and E. E. C. Miranda, "Evaluation of motivational aspects in educational media and technology," in *Proceedings of the 26th World Conference on Educational Media and Technology*, pp. 780–787, AACE, 2014.
- [27] J. N. Marciano, L. C. Miranda, and E. E. C. Miranda, "Applying the method for evaluation of motivational aspects on Karucha ships invaders educational game with brazilian students of Japanese," in *Proceedings of the 20th Workshop de Informática na Escola (WIE '14)*, pp. 66–75, SBC, 2014.

