


Design and Evaluation of a Mobile Application for an Educational Card Game

Rubens Anderson de Sousa Silva*  [UFC - Universidade Federal do Ceará | rubenssilva@great.ufc.br]

Rossana Maria de Castro Andrade†  [UFC - Universidade Federal do Ceará | rossana@ufc.br]

Bosco Borges Aragão Filho  [UFC - Universidade Federal do Ceará | boscoaragao@great.ufc.br]

Ismayle de Sousa Santos  [UFC - Universidade Federal do Ceará | ismaylesantos@great.ufc.br]

Joseane Oliveira Vale Paiva  [UFC - Universidade Federal do Ceará | joseanepaiva@great.ufc.br]

Bruno Sabóia Aragão‡  [UFC - Universidade Federal do Ceará | bruno@great.ufc.br]

Abstract

Educational games can provide players with rich learning and socializing experiences through different interaction paradigms, such as board games, card games, and, more recently, hybrid (physical-digital) games. However, making, maintaining, and evolving an educational analog game is not a trivial task. Balancing the game mechanics and dynamics to provide a pleasant and educational gaming experience can be very difficult to achieve. Furthermore, adding a digital component in the gameplay can disturb the experience of the game and learning objectives, so it is necessary to evaluate this kind of insertion. This work then aims to report the process of developing and evaluating a mobile application for helping the gameplay of a card game that focuses on teaching software testing concepts. Our primary concern during the development of the application was whether its insertion would compromise the learning process or the card game's social experience. The developed application has the following functionalities: point counter, dice scrolling, timer, and a summary of the rules. We designed and developed a first version of the application. Then, we evaluated the impact of its insertion on gameplay by applying the game assisted with the application with students from Computer Science and Computer Engineering courses. By the end of the game, the participants answered questionnaires about the players' experience and their impressions about the application. Based on the results, we perceived that the application provided benefits to the players' experience, although the evaluation highlighted some opportunities for improvement. Thus, we evolve the mobile application based on the comments gathered in this evaluation. This new version has improvements on the user's interface, aiming to provide better user experience and new functionalities. Furthermore, we assessed the second version and compared both versions of the mobile app to collect evidence regarding improvements in the game experience.

Keywords: Hybrid games, Educational games, Board games, Teaching software testing, UX evaluation

1 Introduction

In traditional teaching methods, the focus is on the teacher as responsible for concentrating and organizing the content and passing it to the student, who memorizes and solves activities (Brighenti et al., 2015). Souza and Dourado (2015) argue that this teaching method is less effective for education in the 21st century, because, in the students' perception, the ideal is that he/she is responsible for solving problems and becoming more actively involved with the content exposed in the classroom.

In this sense, the use of educational games can be of great value to aid teaching and improve the learning process. They provide more student engagement, since such games stimulate problem-solving in a practical, playful and, often, collaborative way. For teaching software engineering, as shows Eduardo Battistella and Gresse von Wangenheim (2016), we have a great variety of games, with different level of learning (e.g., skill, cognitive and affective), types (i.e., digital, analog, or hybrid), and subjects, citing as example The GreaTest Card Game (A. Beppe et al., 2018), SimSYS (Longstreet and Cooper, 2011), SCRUMIA (von Wangenheim et al., 2013),

and many others.

In a previous work (A. Beppe et al., 2018), we present the GreaTest Card Game (GTCG). GTCG is an educational game that aims to teach software testing in a playful, practical, and collaborative way for higher education students. After designing the game, we applied sessions of gameplay with classes in the Computer Science and Computer Engineering courses at the Federal University of Ceará, always supervised by a teacher or class monitor. The feedback collected revealed practical difficulties faced by the players during the game related to the game mechanics, and a support application for the management of the matches in progress could mitigate these difficulties.

In this context, we decided to turn the GreaTest Card Game into a hybrid table game, i.e., a game that have digital and physical components relevant to gameplay (Kankainen et al., 2017), by inserting the GreaTest Helper (Borges et al., 2019), a mobile application that aims to assist the management of matches, making some tasks more practical and providing an experience more dynamic to the students. The use of hybrid games can help the teaching, as it brings together the elements and benefits of digital games to the face-to-face and tangible interactivity of elements physicists. GreaTest Helper (GTH) application has three main functionalities: a dice, a point counter, a rule list, and a timer. However, with

*Bolsista da Capes no Programa de Demanda Social

†Bolsista do CNPq de Produtividade em Desenvolvimento Tecnológico e Extensão Inovadora (DT) 2

‡Bolsista da Fundação Cearense de Pesquisa e Cultura.

the insertion of a new element in the gameplay, the concern arose to assess whether the players' experience could be affected.

Aiming to verify if the use of the application exerts any positive or negative influence on the dynamics of the game, we invited a group of undergraduate students of the first semester of Computer Science and Computer Engineering courses to play the game. At the end of the game, the students answered two forms: one, designed by us, with questions about the mobile app, and the other, developed by Petri et al. (2017a), to evaluate aspects of user experience of educational games. The comments of the students collected at this evaluation showed aspects in the mobile app that generate dissatisfaction in the players.

In this paper, we extend our previous work (Borges et al., 2019) by presenting a new version of the mobile application and the results of the new evaluations performed. The improvements in the mobile app were made based on the feedback gathered from the prior evaluation (Borges et al., 2019), improving the interface, functionalities, animations and provided a translated version to English language. Regarding the new assessment, we invited others group of students to play the game using the new version of the GreaTest Helper, and answer the same questionnaires of the previous evaluation. Then, we compared the results of the two versions of the mobile application, and identified positive and negative aspects of the experience related to the mobile application. For instance, the new version brings more dynamicity in the gameplay.

We organized this article as follows: Section 2 presents the theoretical background on the GreaTest Card Game, hybrid games, educational games, software testing, and user experience evaluation in games; Section 3 shows the process of designing and developing the GreaTest Helper application and the gameplay of the GreaTest Card Game using GTH; Section 4 presents the process for evaluating the use of the mobile application; Section 5 details the the results of the evaluation; finally, Section 6 concludes the paper and discusses future works.

2 Background

This study assesses users' perceptions regarding the transformation of an educational card game into a hybrid game through the insertion of a mobile application created to support match management tasks. For this, it is essential to understand some concepts about the GreaTest Card Game itself, educational games, software testing, hybrid games, and about the user experience in games.

2.1 Educational Games

One of the significant challenges in the area of pedagogical teaching is to maintain the focus and motivation of students in academic matters. In the traditional methodology, based on a Jesuit model (Anastasiou, 2001), the student is a mere spectator in lectures focused on the knowledge and reflection of the teacher. However, this methodology has been proved to be less effective, and this can be explained in large part

by the significant changes suffered by both students and society (Brighenti et al., 2015). Students have the perception that solving problems related to the content exposed in the classroom is the most effective way to learn (Brighenti et al., 2015).

Other teaching methodologies seek to center the student on the teaching and learning mechanics, such as the use of the game based methodology (GBL). GBL are defined as the use of game applications for defining learning outcomes (Gresse von Wangenheim and Shull, 2009). GBL approaches apply games with the purpose of learning specific skills and concepts, putting the student in the center of the learning and promotes the resolution of problems through playing a game related to a subject, training or behavior. As explain de Freitas (2006), "the key challenge for effective learning with games is for the learner to be engaged, motivated, supported and interested but also importantly for the learning to be undertaken in relation to clear learning outcomes as well as being made relevant to real world contexts of practice."

2.2 Software Testing

According to Bourque and Fairley (2014), Software testing consists of a dynamic verification of the behavior of a program with a finite set of test cases properly selected from the domain of executions, usually infinite, against the expected behavior. Within the context of the game, the player exercises the understanding of the test type concepts, which are, according to ISO (2013); Legear et al. (2019):

- **Functionality Test:** is determined whether the functional requirements of the test item have been met
- **Acceptance Test:** is performed to assess the system's readiness for deployment and its use by the customer (end-user)
- **Performance test:** aims to assess the performance when the test item faces a "typical" load
- **Stress testing:** aims to assess the performance of the test item when it pushes beyond its anticipated peak load or when available resources (e.g., memory, processor, disk) are reduced below specified minimum requirements, to evaluate how it behaves under extreme conditions
- **Security Test:** evaluates the degree to which a test item and its associated data are protected so that unauthorized persons or systems cannot use, read or modify them and grants only authorized persons or systems required access to them
- **Usability Test:** evaluates whether specified users can use the test item to achieve assigned goals with effectiveness, efficiency, and satisfaction in specified contexts of use

2.3 The GreaTest Card Game

The GreaTest Card Game was developed as an analog card game to assist and put software testing teachings into practice (A. Beppe et al., 2018). Its target audience is composed of students in the area of Computing who want to put knowledge into software testing in a playful way; in special, for students in the beginning of the course.

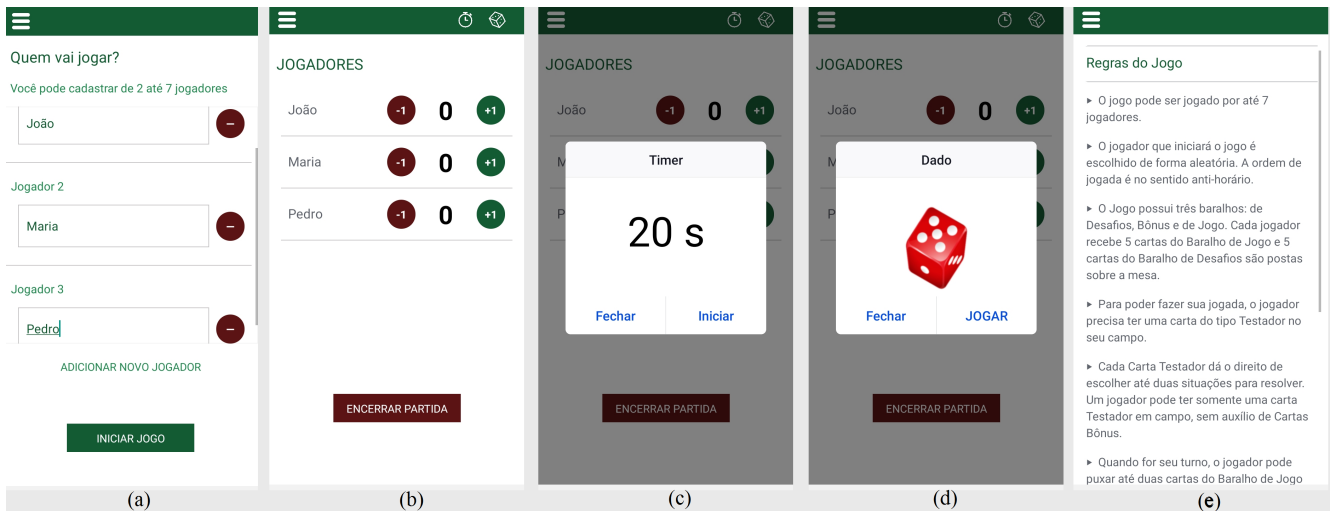


Figure 1. Screenshots from first version of GreaTest Helper Application - Portuguese.

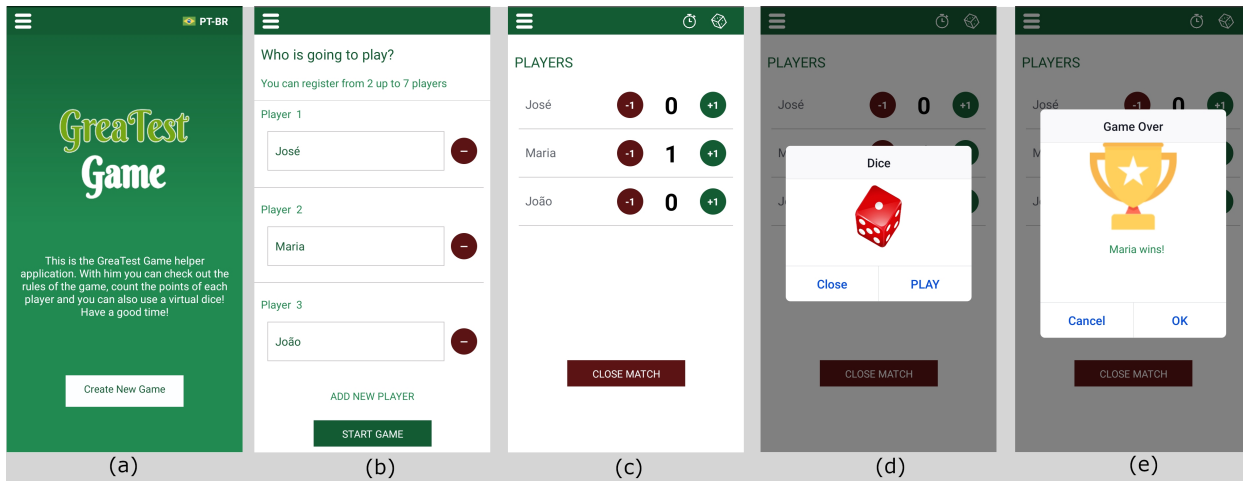


Figure 2. Screenshots from second version of GreaTest Helper Application - English.

The game has three decks: i) Game Deck, which contains the types of tests that players must use to test the scenarios presented (e.g., functional test, safety test), the Tester card, which allows the player performs the tests, and effect cards; ii) Challenge Deck, containing the problem scenarios to be solved by cards from the Game Deck; and iii) Bonus Deck, which contains cards that have direct additional effects on elements of the game (e.g., "trainee", which allows you to do one more test in the same round).

The game is played by three to six players. At the start of the game, each player takes five cards from the Game Deck to the hand, and the players place 5 Challenge Cards on the table, with software problem scenarios that should be tested visible to all players. The gameplay is based on turns, each player can solve up to two problem scenarios, associating the Test Type Cards (e.g., Functional Test) in his hand with the Challenge Cards on the table. However, the player can only play if he has a Tester Card in his field on table. When the player gives the right answer, he can roll the dice and, if the number drawn is present on the back of the answered challenge card, he wins 1 point and draw a card from the Bonus Deck. The player who reaches seven points first wins the game. The access of the game cards and application is

free.¹

2.4 Hybrid Games

The main objective by adopting a hybrid game is to enrich the gaming experience with physical elements to promote new and different types of interaction, e.g., more direct social interaction or with physical objects, ways that would not be possible in exclusively digital games. According to Kankainen et al. (2017), hybrid games are games that combine physical and digital elements in the same product. However, the authors argue that this view can be a limiting factor for its development and analysis, as such games must also communicate conceptual metaphors and hybridity understood as the combination of different cognitive domains that are not usually associated.

Mora et al. (2016) argue that analog games provide engaging social experiences, composed of two levels of interaction: between the players themselves, in person, and also between the players and the physical elements of the game. According to them, this experience is facilitated by tangible, physical, and social interactions, which involve common elements present in various analog games, such as pawns and cards, for

¹Site of the GreaTest Card Game: <http://pesquisa.great.ufc.br/greatest>

example, and also the players themselves. In contrast, the experience of strictly digital games generally makes it difficult to face-to-face interaction. Hybrid table games are also a rich platform for creating and providing new gaming experiences, using functionalities such as touch screens, cameras, and on-line connectivity (Mora et al., 2016).

Xu et al. (2012) argue that digital games tend to have a set of rigid rules, which helps to simplify the programming of the game itself, however blocking possible interactions between players, losing a great potential to provide better experiences. Many digital games have direct means of interaction, e.g., via face-cam, audio stream or chat, but this is very different from being in the same environment as a group of people playing an analog game, being able to negotiate rules, support novice players and lead to interactions beyond the game context. Therefore, physical-digital hybrid games can combine the social experience of analog games with the convenience of digital games, through elements that assist in gameplay, and that is just what we seek with the addition of the GreaTest Helper application in the game GreaTest Card Game.

2.5 Player Experience

User experience (UX), a field of study about human-computer interaction (IHC), has been taking an important place in the discussions fostered, generating a relevant impact on studies developed both in academia and in the industry (Dix, 2017).

Although in recent years, there has been an effort to provide a better understanding of the user experience and to develop a unified view on this area, there is still no consensus on the definition of UX (Law et al., 2008).

Also, there is a wide divergence from what is understood as the quality of user experience, as can be seen in the catalog of user experience assessment instruments developed by Darin et al. (2019). The catalog also shows that there are several applications identified in the cataloged UX assessment instruments, such as mobile devices, hardware and robotics, online platforms, games, and virtual environments. In the context of games, in particular, the user experience is enhanced by recreational and cultural factors and, therefore, it becomes more challenging to be analyzed and evaluated by methods used in other types of interactive systems (Sánchez et al., 2012).

This recreational character and the particular aspects of the user experience in games differentiate them from other traditional types of systems. A particular area studies this given differentiation, the relationship between users and games: Games User Research (GUR). According to Carneiro et al. (2019), GUR is an interdisciplinary field of practice and research that seeks to improve the quality of usability and the experience of users in digital games. Different games can provide different emotional states and arouse different emotional reactions in users, depending on the objectives of development. In this context, educational games have specific objectives related not only to fun and recreational character but also, and primarily, to learning. Considering this type of game as educational material, its evaluation is considered essential for the learning process (Savi et al., 2010). In the evaluation conducted in this study, the questionnaire MEEGA+

(Petri et al., 2017b) assesses the user experience of the constructs addressed by the following qualities: focused attention, fun, challenge, social interaction, confidence, relevance, satisfaction, and usability.

3 GreaTest Card Helper

GreaTest Helper (GTH) is a mobile application developed for the Android operating system, designed to assist in the interaction between the player and the game GreaTest Card Game through the management of player information and the game mechanics itself, giving support and providing the necessary components for the game.

We designed the app so that there is no dependency between it and the game- it is not necessary to use the app to play the game, because: (i) the game was not initially designed to have a digital application as a mechanic, we introduced this feature in another moment with the game completed and evaluated; and (ii) this independence is essential because the player may choose to play the game with analog tools, as real dice, paper, and pen. This section reports the process of designing and developing the GTH application. The app intends to bring together all the crucial features to the game, done in an analog and improvised way.

3.1 Conception Process

Based on the feedback collected during previous application sessions of the game, we noticed that there were practical issues bringing difficulties to the gameplay. The main problems reported by users during the game application sessions were: (1) the game application demanded the use of dice, which was not always available to players; (2) the scoring of points was done in an improvised manner (e.g., paper and pen, notepad), which made the registration of the match an expensive activity and compromised the reliability of the recorded information; (3) sometimes, players took too long to make their move, compromising the progress of the game; and (4) the game has many rules, making the memorization process demand more cognitive effort from the players.

The feedback reported served as the basis for the list of functionalities to be covered by the development of the GTH application. A requirements specification document was created, based on the standard described in Sommerville (2010) for the development of a mobile application. The development process of GreaTest Helper included a (i) phase elicitation, specification, and analysis of requirements, based on the standard described in Sommerville (2010), in which we studied elements commonly used in applications developed to support analog games and how they were adapted, and adopted a product scope based on the demands proposed by the players themselves; (ii) a phase of design, where we choose the framework for development and decisions related to the change in the interfaces were made; (iii) a prototyping phase, where sub-versions of the app were developed and validated by the project members. The development team was composed by one android developer and two members with experience in UX and human-computer interaction to support the choices and application design decisions.

3.2 Functionalities

In this section, we present the GreaTest Helper functionalities, taking into account why each one is necessary, and how they were developed.

3.2.1 Counting points

The point counter functionality helps to solve the problem (2) reported previously. The development needed that only one smartphone was used during the game, in a centralized way, thus preventing constant contact with the device, compromising the players' attention. With that in mind, the layout adopted by the Magic Life Counter 2¹ application seemed appropriate. The application was created in 2014 to help the gameplay of the card game Magic: The Gathering 3², which has the same restriction. This application has a 4.7 rating on Google Play and over five hundred thousand downloads. Figure 3 shows an example of the point counting screen developed for the Magic Life Counter application. As up to 6 players can play the GTCG, the reproduction of total points distribution scheme adopted by the Magic Life Counter was not possible. So, we got inspiration on the way the points increase and decrease for the players. They can enter their names via keyboard on the Registration Screen (Figure 2b), and the point counters are automatically associated with the names on the Game Screen. Then, the player presses the +1 or -1 button to increase or decrease one point, respectively. Figure 2c illustrates the implementation of this feature in the GTH.



Figure 3. Screenshot of the app Magic Life Counter

3.2.2 Dice

The Dice functionality supports players if they do not have real dice to play. Following the model of a physical dice, the App draws random numbers from 1 up to 6, with equal probability for each result, and the interaction uses the interface metaphor (Jennifer Preece, 2013) of a real dice scrolling action. The function can be accessed quickly from the Game Screen of the application (see Figure 2c) by clicking on the icon that refers to the image of a dice (affordance).

3.2.3 Timer

In the evaluation of the GTCG (A. Beppe et al., 2018), a user suggested inserting a one-minute timer to limit each player's playing time during the turn. Once the time is up, the player in question would be obliged to perform an action or skip the turn. As the GTCG is an educational game, it is necessary to ensure that any change in the game, such as changes in rules, dynamics, or insertion of digital components, does not compromise the main objective of the game, which is to improve the students' learning process Dempsey et al. (1996).

The act of limiting the playing time would compromise the student's immersion process within the problem established there. Besides, the need to interact with the application whenever a player starts and finishes a move could negatively impact the duration of matches and the natural progress of the game. Taking this into consideration, we decided to insert a timer as an optional tool, which can be triggered if the other players agree that the current player is exceeding the duration of that move. The duration of the timer is 20 seconds (see Figure 1c.), leaving players to decide whether or not to renew the count. For the access to this functionality, the same standard used for the dice was adopted, ensuring that the option is always available throughout the game.

3.2.4 Game Rules

The Game Rules functionality, which the player can access via the menu, is composed only of textual elements that provide a summary of the GTCG manual, making it easier to consult the rules if the players need - problem (4) reported in the previous section. Figure 1e shows the current state of the developed functionality.

3.3 Gameplay using the App

Before starting the game, it is necessary to know who starts playing, and for that, the players can use the application dice (Figure 2d) to draw a number for each player, and the player who gets the highest number starts the game. Given the start of the game, the player begins to make his move (as described in Section 2). If delayed, players can agree to use the 20-second timer present in the GTH to indicate that time to play is limited. If the player succeeds and sets the correct type of test for the scenario, he must again use the dice present in the GTH to draw a number and check if he wins a point. If so, he must press the +1 button next to his name (Figure 2c) to increase a point. If he reaches 7 points, the match can end by pressing the *END MATCH* button and a pop-up will appear with the winner's name. At any time during the game, if any

¹App Magic Life Counter: <https://talisman.games/products/magic-life-total/>

²Site of the game Magic: The Gathering 3: <https://magic.wizards.com/pt-br>

player has doubts about the rules of the game, he can consult them by accessing the menu and opening the Rules Screen, which are not yet translated to English language.

3.4 App Improvements

Based on the feedback collected from the evaluation performed in our previous work (Borges et al., 2019), we decided that the GreaTest Helper App needs improvements, especially involving user experience feedback. From the students' responses, we identified the need for improving the dice functionality, which should have a more striking animation. In the first version of GTH, the dice just changes the number result. Then, for improve the dice, we implemented an animation of 1 second for clearly show that there was a dice roll, in which the dice grows to 1.3 times its size and presents bounce animation, shows the number result and returns to the original size; and plays a noise of rolling dice while the animation happens.

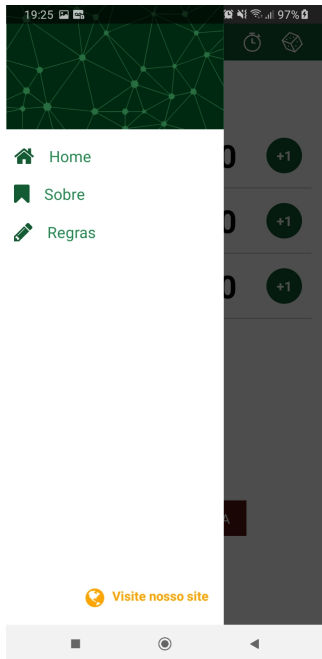


Figure 4. Menu from GreaTest Helper - Second Version

Besides that, we observed a Nielsen (1994)'s usability issue on preventing error by the user: when using the App, the user could accidentally press the physical or digital button to exit the screen, losing the context created until then, e.g., being in the Game Configuration Screen (Figure 2b) and leaving unintentionally, losing the students' names written until then. We decided that whenever there is volatile information on the screen, the GTH should show confirmation mechanisms via pop-ups through which the user must confirm that he wants to leave (see Figure 5).

We have improved the responsiveness of the App; the screen shows similarly at different screen resolutions on different devices. Also, we improved the aesthetic details of the options menu and made the link to our website, providing access to information about the game, publications, and developers (Figure 4).

Finally, we added the translation of the App to the English language; now, the player just have to press the button on

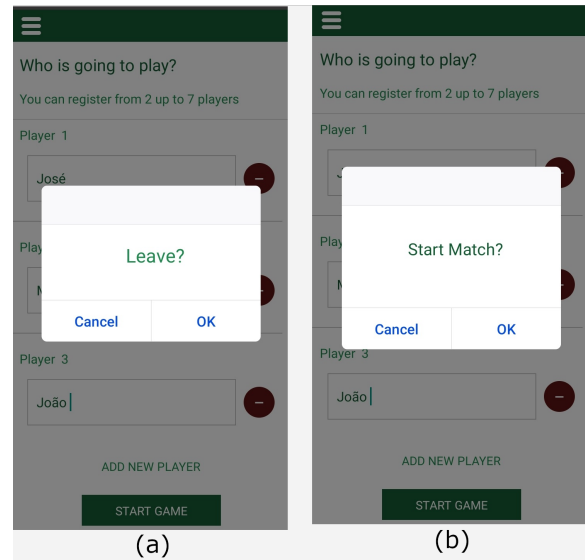


Figure 5. Confirmation pop-up - Second version.

the home screen (see Figure 2a) to change the language from Portuguese to English and vice-versa.

A comparison between the general interfaces of the two versions are shown in Figure 2 and 1. Also, other differences on the interfaces are shown on Figure 5a and b, which the previous version did not have them, and on the dice functionality, but as it is an animation, can not be illustrated on Images.

4 Evaluation

The purpose of evaluating the GreaTest Helper application was to verify if the application influences the GreaTest Card Game gameplay and, in this case, to identify if the influence is positive or negative, seeking to detect how the application helped or hindered the progress of the game. Also, the evaluation aimed to verify aspects of the player experience with the use of the mobile app, from the users' point of view, considering aspects such as focused attention, fun, challenge, social interaction, trust, relevance, satisfaction, and usability (Petri et al., 2017b). For this purpose, we applied two questionnaires to the participants of the evaluation after the gameplay.

4.1 The questionnaires

In order to assess aspects of player experience, we used the MEEGA+ questionnaire (Petri et al., 2017b), which intends to the evaluation of educational games, whether analog or digital. The version for digital games includes three questions about protection from errors, which were not used in this work because the evaluation is about a game which has an analog core, even using a digital application. The division of the MEEGA+ questionnaire into three sections is as follows: the first addresses questions about the user's profile, such as age, gender, and frequency with which the user usually plays digital and analog games; the second section deals with questions about aspects of the player experience and the third consists of questions about the perception of learning.

The mobile app was useful for the game. *

1 2 3 4 5

Strongly disagree Strongly agree

The mobile app was essential for the game. *

1 2 3 4 5

Strongly disagree Strongly agree

The mobile app made the game more dynamic. *

1 2 3 4 5

Strongly disagree Strongly agree

The mobile app generated difficulties in the progress of the game. *

1 2 3 4 5

Strongly disagree Strongly agree

The mobile app slowed down the game. *

1 2 3 4 5

Strongly disagree Strongly agree

Figure 6. Questions from Questionnaire specific to GTH.

Although this last topic is not part of the present research scope, we kept it to maintain the integrity of the application of the questionnaire.

Besides, to evaluate specific aspects of the influence of the GHT app in the gameplay and the game progress, we applied an online semi-open questionnaire that contained specific questions about the use of the application. The objective and closed questions used a five-point Likert scale, which varied from "strongly disagree" to "strongly agree" - as shown in Figure 6; and an open and optional question in which students could make some additional considerations about using the mobile app. The questions in this questionnaire investigated whether, in the students' perception, GTH: (1) was useful for the game; (2) it was essential to the game; (3) made the game more dynamic; (4) generated difficulties in the progress of the game; and (5) slowed down the game. For both questionnaires, the responses were anonymous, and the students had autonomy and privacy to answer them.

4.2 First Evaluation

The first evaluation used the card game and the first version of the GreaTest Helper (Borges et al., 2019). The participants were ten students from the first semester of the Computer Engineering and Computer Science courses at the Federal University of Ceará, who volunteered to participate in the study. Seven (70%) out of ten participants of the evaluation were

male, and three (30%) were female, all of them were under 28 years old and were in the first semester of their courses. Out of the ten users, four (40%) stated that they rarely play digital games, one (10%) declared that he played at least once a month, three (30%) stated that they played weekly, and two (20%) played daily.

As the GreaTest Card Game deals with the teaching of Software Testing and the students of the first semester had not yet seen subjects that substantially addressed this topic, we developed an introductory software testing material and explained to the students. The material consisted of a slide show that introduced the meaning of software testing, its importance in the software development process, some of the most common types of testing, and the roles and functions of the professionals involved in the testing activities. Then, after explaining the material, we explained the rules of the game, the types of cards and their functions, and we gave the students a printed copy of the game manual; after that, students received a smartphone running the first version of the GTH application and were informed about its functions (dice, timer, point counter and summary of the rules).

With the manual and the application in hand, the first round of the game began, as can be seen in Figure 7. During the entire application of the game, we were present to clarify any doubts or questions.

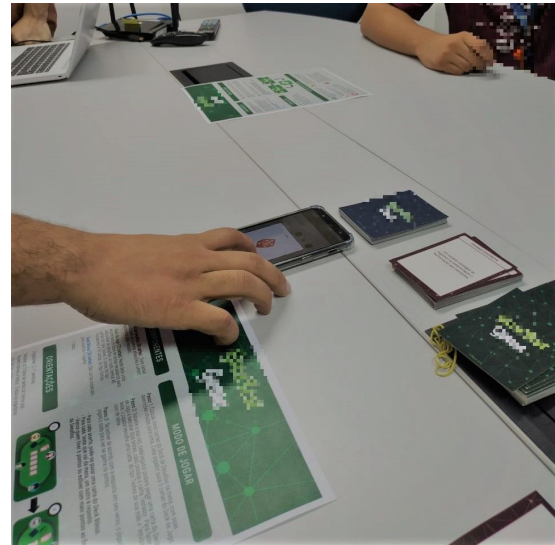


Figure 7. First Game session of the GreaTest Card Game. (Borges et al., 2019)

After playing the GTCG, students answered the two questionnaires that addressed the objectives of the evaluation (assessing aspects of the player experience and the students' perception about the influence of the GTH app within the game).

4.3 Second Evaluation

After the first evaluation, using the first version of the GreaTest Helper, we collected and analyzed the data of the questionnaires. Its results showed some aspects to fix on the app that caused impairments to the player experience with the game. Based on these aspects, detailed in section App Improvements, we developed a new version of the app, improving some functionalities. Then we ran the second evaluation,



Figure 8. Second Game session from GreaTest Card Game.

repeating the same process as First Evaluation, but using the new version of the GTH and with other students with the same profile. After playing the game, the participants answered the same two questionnaires about the app and the player experience. Figure 8 illustrates the application of the game.

In the second evaluation, there were eleven participants, ten (90%) male, and one (10%) female. They were all under 28 years old and students of Computer Science and Computer Engineering courses, just like the first evaluation. Out of the eleven participants, four (40%) stated that played digital games daily, three (30%) declared that played weekly, one (10%) played at least once per month and also one (10%) played just rarely - one (10%) of the participants of the second evaluation did not answer the question about how often the participants played digital games. The data collected in the questionnaires showed some positive and negative results, which we discuss in the following section.

5 Results

In this section, we present the results obtained from the data gathered in the questionnaires applied in the two evaluation process.

5.1 First Evaluation Results

In the first evaluation process, we applied the two questionnaires (post-game questionnaire and MEEGA+) after the participants play the game with the first version of the app GreaTest Helper.

5.1.1 Post-game questionnaire results

The post-game questionnaire consists of specific questions about the influence of the application in GTCG gameplay. In the first evaluation, eight out of the ten participants answered this questionnaire after the game was applied. The objective questions (answered with a Likert scale from 1 to 5) dealt with how much the application influenced the dynamics and speed of the game, how useful and vital it was for the game, and how much the application facilitated the game’s progress.

In the question about how much the application made the game more dynamic, seven out of the eight students (87.5%) answered with 4 or 5, and one (12.5%) answered 3 so that the average was 4.38.

Regarding how much students found the app useful for the game, all the eight responses varied between 4 and 5, and the average was 4.75. In the statement that the application was essential for the game, the answers varied more (from 2 to 5), and the average for this question was 3.25. In the question that stated that the application generated difficulties in the flow of the game, seven of the eight responses (87.5%) ranged between 2 and 1, one student answered 4, and the average was 1.63. In the last objective question, which deals with how much the application slowed down the game, the responses ranged from 1 to 3, and the average was equal to 1.75. We show all the questions with their means and medians in Figure 9.

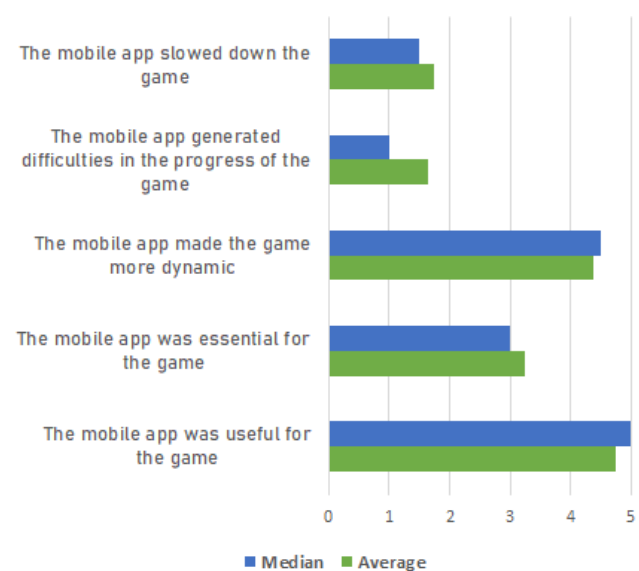


Figure 9. Averages and medians of post-game questionnaire responses in the first evaluation.

Also, the questionnaire had an open question - not a mandatory one - so that students could report any additional comments about the app. Four out of the eight respondents commented on this question, and they talk about animating the image of the dice in its functionality. Users reported that the animation of the dice thrown was very subtle, to the point of being imperceptible if there was a change when the dice results in a number equal to the previous one (e.g., when the dice was already at number 4 and, after being thrown, falls off again at number 4). One of the players also reported that he felt impaired during the game due to this problem.

5.1.2 MEEGA+ Results

The MEEGA + questionnaire was applied together with the post-game questionnaire, in order to evaluate aspects of player experience, which are: focused attention, fun, challenge, social interaction, trust, relevance, satisfaction and usability (subdivided in aesthetics, learnability, operability, and accessibility). The ten students who played the game

in the first evaluation answered the MEEGA+ questionnaire. For each of the user experience qualities addressed by the questionnaire, there were two to four objective questions (answered with a Likert scale) - excepting usability, which had nine questions, all also objective. For the analysis of MEEGA+, we calculated the average and median of each aspect of user experience, which consisted of the arithmetic mean of the average of the questions referring to each specific aspect, as can be seen in Figure 10. The appendix shows all the questions of the MEEGA+ questionnaire and their responses.

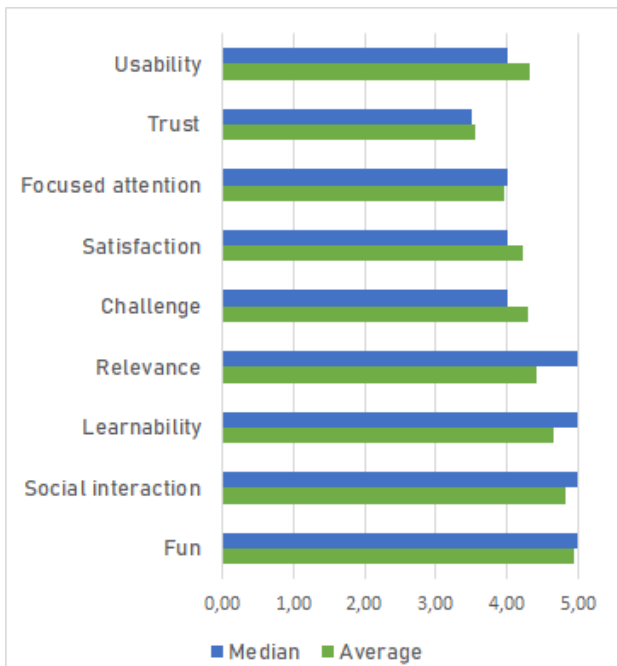


Figure 10. Averages and medians of UX components from MEEGA+ Questionnaire in the first evaluation.

Within the aspect of focused attention, the questions dealt with something that caught the user’s attention at the beginning of the game, the user’s involvement, and how it affected their sense of time and space while playing. The average of this aspect was 3.97.

The questions approaching the fun of the game dealt with how the user had fun with the game. The average of questions of this quality was 4.95, with the Player Experience aspect having the highest average.

The challenging aspect had questions about the game being adequately challenging, offering new challenges at an appropriate pace, and not becoming monotonous in their tasks. The average of this aspect was 4.30.

Approaching social interaction, the questions concerned how much the user can interact with other people during the game, the promotion of moments of cooperation and competition for the game, and also about how well the user felt interacting with other people during the game. This aspect had an average of 4.83.

The quality of user experience that deals with confidence had questions about the impression of ease of the game at first sight of the users and about the organization of the content, giving them confidence that they would learn from the game. The average of this quality was the lowest among the aspects

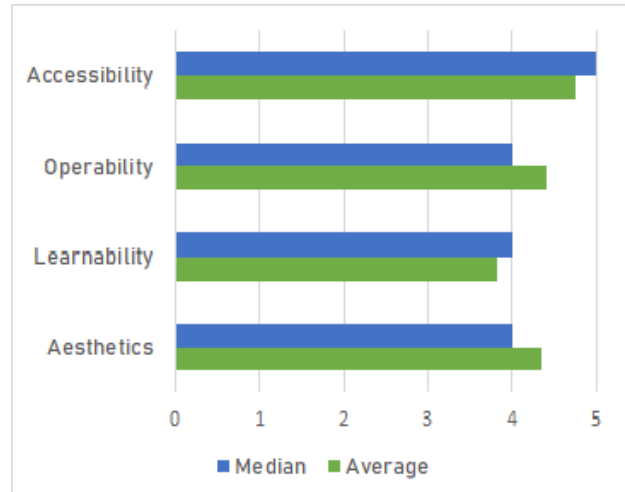


Figure 11. Averages and medians of Usability components from MEEGA+ Questionnaire in the first evaluation.

of UX evaluated, being equal to 3.55.

The relevance questions approach how relevant the game’s content was to the user’s interests, how much the users perceived game’s content was related to the subjects of their course, the suitability of the game as a teaching method for the subject in question and also how much the user preferred to learn from the game than in other ways. The average of responses in this regard was 4.43.

The satisfaction questions approach how satisfied the user was by completing game tasks and with things he learned in the game, the possibility of him recommending the game to colleagues, and also how he realized he was progressing in the game due to his effort. The average of this aspect of UX was 4.23.

The aspect of learning, also related to the user experience, had questions about how users prefer to learn with the game than with other methods, and how efficient the game was for the learning. The average of this aspect was 4.67.

The questions about aesthetics, learnability, operability, and accessibility divide the usability aspect. The average of the usability question was 4.25. Figure 11 shows the averages for each sub-aspect of usability.

Also, MEEGA + had two open questions for students to report the strengths of the game and another to suggest improvements. The content of the answers to both questions was categorized and counted. The most recurring themes in the game’s strengths, reported by the students, were about the learning acquired through the game, the fun that the game provides and compliments about rules and mechanics; these themes appeared five, four and three times, respectively - so that each response could address several themes. We identified in the suggestions for improvement only three categories of responses. From the responses of the ten students, five dealt with improvements in the application, and six also suggested changes and improvements in the rules and mechanics of the game, while one response suggested adding educational content to some cards. Out of the five responses which suggested improvements to the mobile application, three suggested changes in the dice animation, such as ”it is necessary to analyze the possibility an animation or something that sim-

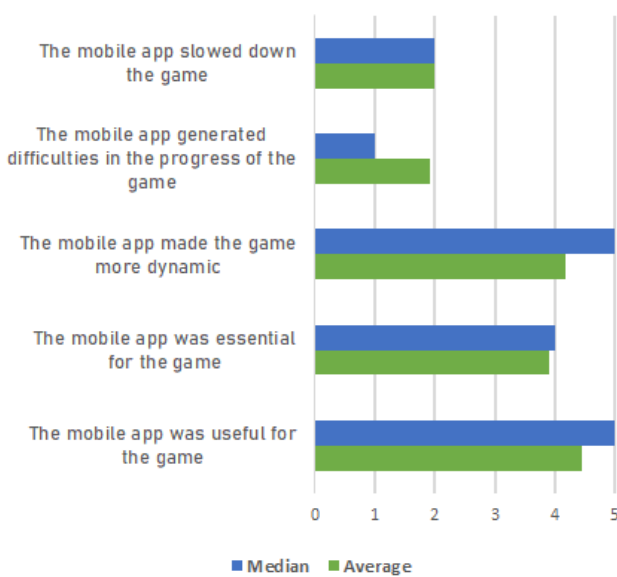


Figure 12. Averages and medians of post-game questionnaire responses in the second evaluation.

ulates the throw of dice. (freely translated from Portuguese)”.

5.2 Second Evaluation Results

In the second evaluation process, we applied the same process, with the same two questionnaires, to evaluate aspects of player experience and the influence of the new version of GTH app in the gameplay.

5.2.1 Post-game questionnaire Results

All the eleven participants of the second evaluation answered the post-game questionnaire. In the first question, about how much the application made the game more dynamic, the answers showed more varied opinions than the first evaluation, so that seven of the eleven students answered with 5, three answered 3.0, and one answered 2.0, resulting in an average of 4.18.

Regarding the perception of the participants about the app useful to play the game, nine responses varied between 4 and 5, and two participants answered 3; so the average was 4.45. In the third question, about the application to be essential for the game, seven participants answered 4 or 5, two answered 3, and two answered 2, so the average for this question was 3.91, which was lower than in the first evaluation. In the statement that the application generated difficulties in the flow of the game, eight out of the ten responses ranged between 2 and 1, two students answered 4 or 5, and one participant answered 3, so the average was 1.91. In the fifth question, about how much the application slowed down the game, the responses ranged from 1 to 3, as 8 participants answered 1 or 2, and three answered 3, so the average was 2.00. We show all the questions with their averages and medians in Figure 12.

In the optional open question of the post-game questionnaire, nine of the eleven participants commented. Three out of nine answers were about the confusion generated by the dice animation, even with the corrections made in the new version. One of the respondents commented eulogizing the

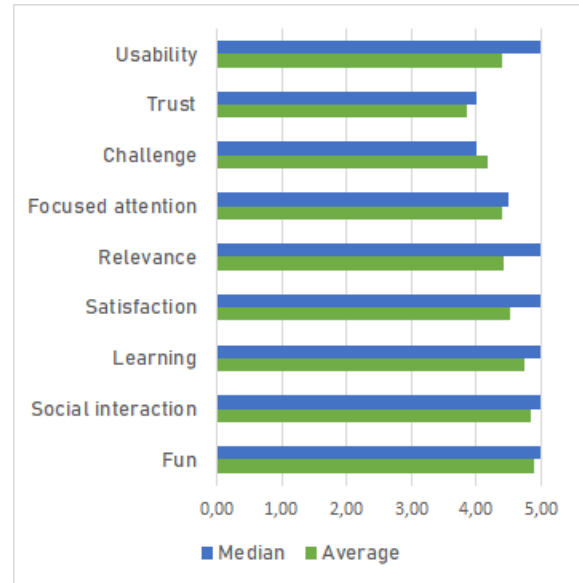


Figure 13. Averages and medians of UX components from MEEGA+ Questionnaire in the second evaluation.

efficiency and reliability of the point-counter of the application. Another one participant commented about the app depriving the players of the experience of playing with physical dice. The four other answers were about possible changes in the game mechanics and rules; however, this was not the focus of this evaluation.

5.2.2 MEEGA+ Results

Ten out of the eleven participants of the second evaluation process answered the MEEGA+ questionnaire. As in the first evaluation, we analyzed the data gathered in the MEEGA+ calculating the average and median of each aspect of user experience, which consisted of the arithmetic mean of the average of the questions referring to each specific aspect, as can be seen in Figure 13. The appendix shows all the questions of the MEEGA+ questionnaire and their responses in the second evaluation.

The average of the responses for the questions about fun was 4.90, which was the higher mean of the aspects of player experience, followed by social interaction (average 4.83). Even if the goal of the research was not evaluate learning, we applied the questions about this aspect contained in MEEGA+ to maintain the validation of the instrument, and the aspect of learning had the third higher average, which was 4.73. The average of the aspect of satisfaction was 4.53, followed by relevance, 4.43, and focused attention, 4.40. The Challenge aspect, which had an average under 4 in the first evaluation, showed a higher value in the second evaluation, which was 4.17. The aspects of confidence had the lowest average with 3.85.

The average of the questions of the aspects of usability (aesthetics, learnability, operability, and accessibility) showed an average of 4.40. Figure 14 shows the averages for each sub-aspect of usability.

In the MEEGA+ open questions, all of the ten respondents of this questionnaire commented about advantages and suggestions for improvement. The answers to these questions were categorized and counted. In the question about the strengths of the game, the four most recurrent subjects com-

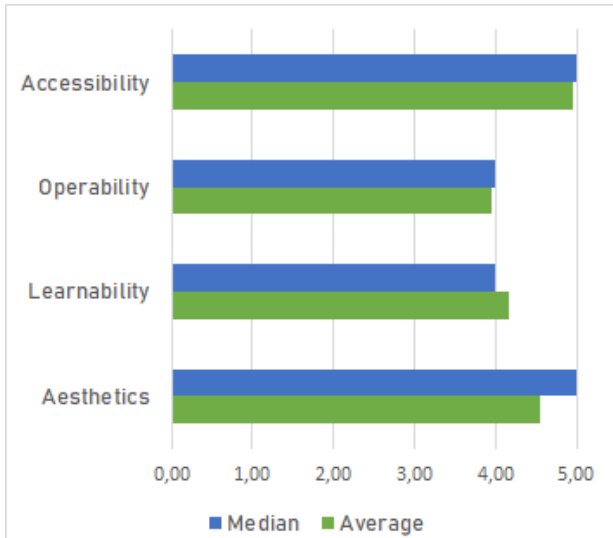


Figure 14. Averages and medians of Usability components from MEEGA+ Questionnaire in the second evaluation.

mented were about: (i) the learning with the game, which appeared in six answers; (ii) the interaction with other players, which occurred in four comments; (iii) the competitiveness of the game, and (iv) the dynamism, both appeared in three answers.

The question about the weaknesses of the game has eight answers, which we categorized into three groups. The most recurrent category had comments about changes in the rules and mechanics of the game, which appeared in all of the eight answers. The other two categories appeared only in one comment, each one, and they were comments about the inclusion of digital elements in the game and the content of the cards.

5.3 Comparisons

We compared the results collected in the first and the second evaluation process, aiming to identify relevant differences in what the participants reported about the different versions of the mobile app.

The data gathered by the post-game questionnaire in the two evaluations revealed that the first version of the app showed slightly higher averages in the responses about how much the GreaTest Helper made the game more dynamic and how useful it was to the gameplay. In contrast, the second version showed better results in the questions about how essential the app was to play the game, whether the application hampered the game flow and how the app slowed down the gameplay, as shown in Figure 15.

Furthermore, the second version of the mobile app showed better medians for all questions - except for the question about the app hampered the flow of the game, which already had a median of one in the first evaluation and remained so in the second evaluation process, and the question about the usefulness of the mobile app, which had a median of 5 in both evaluations. Figure 16 shows the comparisons of medians from the two applications of the post-game questionnaire.

In the subjective question of the post-game questionnaire, in which the participants made additional comments about the app GreaTest Helper, the first evaluation had four an-

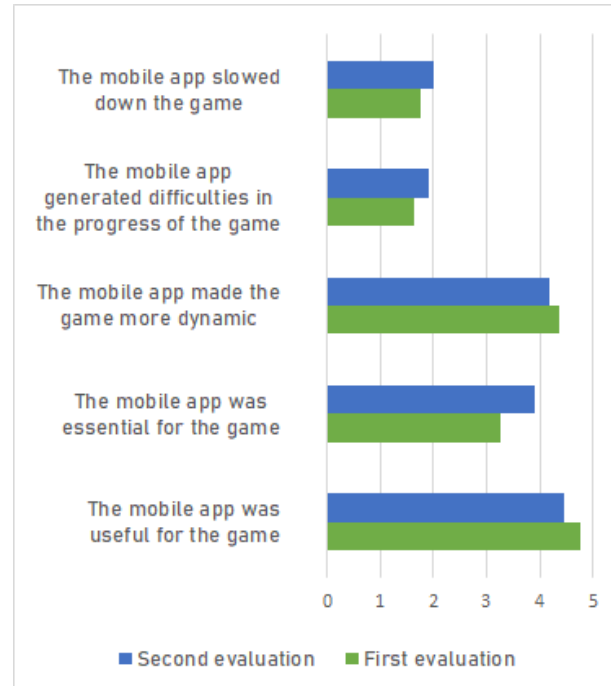


Figure 15. Comparison of the averages of the questions from the post-game questionnaire

swers, and all of them were about the lack of animation of the dice feature, that generated some confusion during the gameplay. On the other hand, the answers in the second evaluation, with the new version of the GreaTest Helper, showed four categories of comments: (i) changes in the mechanics and rules of the game; (ii) about the app efficacy; (iii) the lack of experience with physical elements of the game (such as dice) - provided by the application; and (iv) about the lack of clarity of the dice animation.

6 Conclusions

This study presented the hybridization process of an educational analog card game for teaching software testing in ludic, practical and collaborative way, called GreaTest Card Game (GTCG). This new version of the game inserted the digital application GreaTest Helper (GTH) in the gameplay as a mechanic and supporting tool. Since the insertion of such tool could arise concerns on gameplay, experience, and learning process, we evaluated the first version of the GTH application and, based on the results, we improved the app. Next, we evaluated the second version of GTH and compared the results of both versions.

The two evaluations showed very similar results in the aspects of the player experience evaluated by the MEEGA+ questionnaire, but the second evaluation revealed slightly better results. Furthermore, in the aspects of usability (aesthetics, learnability, operability, and accessibility), the second version of the application performed better.

It is important to notice that in both evaluations, the results obtained from the MEEGA+ questionnaire revealed good averages in several aspects of the user experience, mainly about the fun provided by the game - allied with the GTH app - and also to the social interaction and learning with the game that the students reported experiencing. Despite this, the aspects

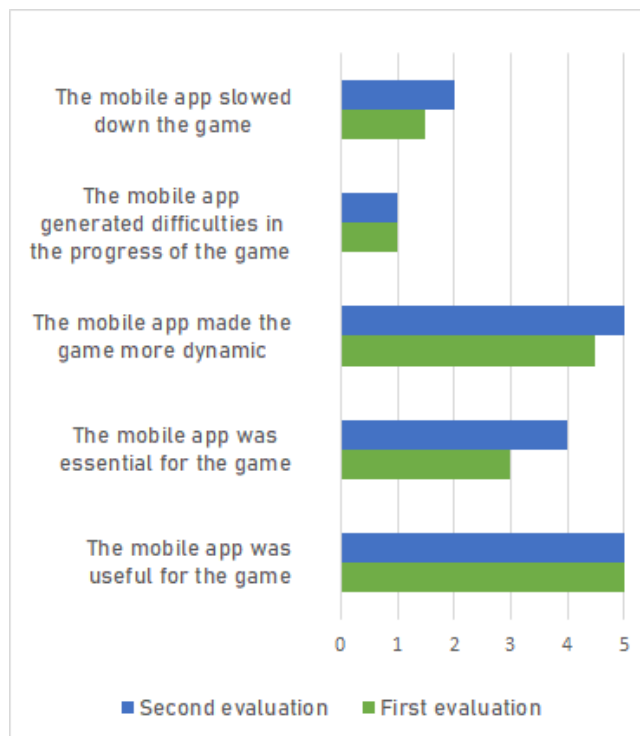


Figure 16. Comparison of the medians of the questions from the post-game questionnaire

of trust and focused attention performed worse. This may be related with the constant interaction with an device aside the game, and the issues about the app's functionalities.

According to the data obtained and analyzed from the post-game questionnaire about the influence of the two versions of the GreaTest Helper application in the GreaTest Card Game, it is notable that the application was useful for the progress of the game, even if it is not possible to say whether it is essential for the application of the GTCG. Furthermore, the students' responses revealed that the application did not slow down the game or generate any relevant difficulties in the game flow.

Despite the students' positive perception of the application's usefulness, still, improvements can be designed and executed in some of its functionalities. The purpose of the second version of the app was to fix some problems identified in the first evaluation process, as the problematic animation of the dice, which makes it difficult to perceive that the button that triggers the dice roll responds correctly to the command. However, users of the second app version still commented about the confusion generated by the animation of the dice. They also commented about crucial aspects of the experience with the GTCG and the GTH, such as the lack of experience with physical elements (which are an inherent part of analog games). These points reported by the users and the results extracted from the data analysis reveal that there is still a need for improvements in the GreaTest Helper application and also in its integration with the GreaTest Card Game, including to make the game more hybrid.

Furthermore, we can see that there is a lack of studies that specifically propose methods of evaluating hybrid games and this ended up causing difficulties in ensuring rigor when using the MEEGA+ form to evaluate a game that has some level of hybridity. However, we tried to overcome this dif-

ficulty by inserting a set of specific questions to evaluate the application, trying to understand, from the students' perspective, the impact of using GreaTest Helper on gameplay.

The analysis of the data collected through the questionnaires, including direct feedback from users, revealed possibilities for updates and improvements in both the GreaTest card game and the GreaTest Helper application. As future work, we intend to develop a new version of the game with reviewed and new mechanics integrating GTCG, and investigating new possibilities of hybridization and pervasiveness -however, holding the option to play without it.

References

- (2013). *ISO/IEC/IEEE P29119-4-DISMay2013*. IEEE Software and Systems Engineering Standards Committee.
- A. Beppe, T. Linhares de Araújo, I. Aragão, B., Santos, I. D. S., Ximenes, D., and M. Castro Andrade, R. (2018). Greatest: a card game to motivate the software testing learning. pages 298–307.
- Anastasiou, L. G. C. (2001). *Teaching Methodology at the Brazilian University: elements of a trajectory. (In portuguese, Metodologia de Ensino na Universidade Brasileira: elementos de uma trajetória.)*. Papirus.
- Borges, B., Silva, R., Paiva, J., Aragão, B., Santos, I., and Andrade, R. (2019). Design e avaliação de um aplicativo móvel complementar para um jogo de cartas educacional. In *Anais do I Workshop sobre Interação e Pesquisa de Usuários no Desenvolvimento de Jogos*, pages 21–30, Porto Alegre, RS, Brasil. SBC.
- Bourque, P. and Fairley, R. (2014). *Guide to the Software Engineering Body of Knowledge*, volume 3. IEEE Computer Society.
- Brighenti, J., Biavatti, V., and Rodrigues de Souza, T. (2015). Teaching-learning methodologies: an approach under the students' perception. (in portuguese, metodologias de ensino-aprendizagem: uma abordagem sob a percepção dos alunos). *Revista Gestão Universitária na América Latina : Revista GUAL*, 8.
- Carneiro, N., Darin, T., and Viana, W. (2019). Application of games user research to the evaluation of games based on location. (in portuguese, Análise da Aplicação de Games User Research à Avaliação de Jogos Baseados em Localização).
- Darin, T., Coelho, B., and Borges, B. (2019). Which Instrument Should I Use? Supporting Decision-Making About the Evaluation of User Experience. pages 49–67.
- de Freitas, S. (2006). Learning in immersive worlds: a review of game-based learning. Workingpaper.
- Dempsey, J., Lucassen, B., and Rasmussen, K. (1996). *The Instructional Gaming Literature: Implications and 99 Sources*. University of South Carolina, College of Education.
- Dix, A. (2017). Human-computer interaction, foundations and new paradigms. *Journal of Visual Languages and Computing*.
- Eduardo Battistella, P. and Gresse von Wangenheim, C. (2016). Games for teaching computing in higher education

- a systematic review. *IEEE Technology and Engineering Education*, 1:8–30.
- Gresse von Wangenheim, C. and Shull, F. (2009). To game or not to game? *Software, IEEE*, 26:92–94.
- Jennifer Preece, Helen Sharp, Y. R. (2013). *Interaction Design: Beyond Human-Computer Interaction*. (In portuguese, *Design de interação: Além da Interação Humano-Computador*). Bookman, Porto Alegre, Brasil, 3 edition.
- Kankainen, V., Arjoranta, J., and Nummenmaa, T. (2017). Games as blends : Understanding hybrid games. *Journal of Virtual Reality and Broadcasting*.
- Law, L.-C., Roto, V., Vermeeren, A., Kort, J., and Hassenzahl, M. (2008). Towards a shared definition of user experience. pages 2395–2398.
- Legard, B., Denoo, O., Friedenberg, D., Kramer, A., and Zmitrowicz, K. (2019). *Foundation Level Syllabus – Acceptance Testing*. International Qualification Board for Business Analysis - ISTQB.
- Longstreet, C. S. and Cooper, K. (2011). Using games in software engineering education to increase student success and retention. page 554.
- Mora, S., Loreto, I. D., and Divitini, M. (2016). From interactive surfaces to interactive game pieces in hybrid board games. *JAISE*, 8:531–548.
- Nielsen, J. (1994). Usability inspection methods. In *Conference companion on Human factors in computing systems*, pages 413–414.
- Petri, G., Gresse von Wangenheim, C., and Borgatto, A. (2017a). Evolution of a game evaluation model for the teaching of computing. (in portuguese, evolução de um modelo de avaliação de jogos para o ensino de computação).
- Petri, G., Gresse von Wangenheim, C., and Borgatto, A. F. (2017b). *MEEGA+, Systematic Model to Evaluate Educational Games*, pages 1–7. Springer International Publishing, Cham.
- Savi, R., Gresse von Wangenheim, C., Ulbricht, V., and Vanzin, T. (2010). Proposal for an educational games evaluation model. (in portuguese, proposta de um modelo de avaliação de jogos educacionais). *RENOTE*, 8.
- Sommerville, I. (2010). *Software Engineering*. Addison-Wesley Publishing Company, USA, 9th edition.
- Souza, S. C. d. and Dourado, L. (2015). Problem-based learning (pbl): an innovative learning method for educational teaching (in portuguese, aprendizagem baseada em problemas (abp): um método de aprendizagem inovador para o ensino educativo). *Holos*, 5:182–200.
- Sánchez, J. L. G., Vela, F. L. G., Simarro, F. M., and Padilla-Zea, N. (2012). Playability: analysing user experience in video games. *Behaviour & Information Technology*, 31(10):1033–1054.
- von Wangenheim, C. G., Savi, R., and Borgatto, A. F. (2013). Scrumia—an educational game for teaching scrum in computing courses. *Journal of Systems and Software*, 86(10):2675 – 2687.
- Xu, Y., Barba, E., Radu, I., Gandy, M., and Macintyre, B. (2012). Chores are fun: Understanding social play in board games for digital tabletop game design. *Proceedings of DiGRA 2011 Conference: Think Design Play*.

APPENDIX

Table 1. Part I - Questions and answers regarding the player experience in the first evaluation with the MEEGA+ questionnaire.














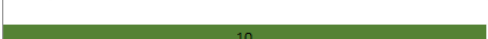










		USER EXPERIENCE	MEDIAN
Trust	When I first looked at the game, I had the impression that it would be easy for me.		3
	The contents and structure helped me to become confident that I would learn with this game		4
Challenge	This game is appropriately challenging for me.		4
	The game provides new challenges (offers new obstacles, situations or variations) at an appropriate pace		4
	The game does not become monotonous as it progresses (repetitive or boring tasks).		4
Satisfaction	Completing the game tasks gave me a satisfying feeling of accomplishment.		4
	It is due to my personal effort that I managed to advance in the game.		4
	I feel satisfied with the things that I learned from the game.		4
	I would recommend this game to my colleagues.		5
Social interaction	I was able to interact with other players during the game.		5
	The game promotes cooperation and/or competition among the players.		5
	I felt good interacting with other players during the game.		5
Fun	I had fun with the game		5
	Something happened during the game (game elements, competition, etc.) which made me smile.		5
Focused attention	There was something interesting at the beginning of the game that captured my attention		3
	I was so involved in my gaming task that I lost track of time.		4
	I forgot about my immediate surroundings while playing this game.		4
Relevance	The game contents are relevant to my interests.		4
	It is clear to me how the contents of the game are related to the course.		5
	This game is an adequate teaching method for this course.		4
Perception of learning	I prefer learning with this game to learning through other ways (e.g. other teaching methods).		4
	The game contributed to my learning in this course		5
	The game allowed for efficient learning compared with other activities in the course.		5
	The game contributed to teach software testing.		5

Table 2. Part II - Questions and answers regarding the player experience in the second evaluation with the MEEGA+ questionnaire.

		USER EXPERIENCE	MEDIAN
Trust	When I first looked at the game, I had the impression that it would be easy for me.		3
	The contents and structure helped me to become confident that I would learn with this game		4
Challenge	This game is appropriately challenging for me.		5
	The game provides new challenges (offers new obstacles, situations or variations) at an appropriate pace		5
	The game does not become monotonous as it progresses (repetitive or boring tasks).		4
Satisfaction	Completing the game tasks gave me a satisfying feeling of accomplishment.		5
	It is due to my personal effort that I managed to advance in the game.		4
	I feel satisfied with the things that I learned from the game.		5
	I would recommend this game to my colleagues.		5
Social interaction	I was able to interact with other players during the game.		5
	The game promotes cooperation and/or competition among the players.		5
	I felt good interacting with other players during the game.		5
Fun	I had fun with the game		5
	Something happened during the game (game elements, competition, etc.) which made me smile.		5
Focused attention	There was something interesting at the beginning of the game that captured my attention		4
	I was so involved in my gaming task that I lost track of time.		5
	I forgot about my immediate surroundings while playing this game.		4
Relevance	The game contents are relevant to my interests.		5
	It is clear to me how the contents of the game are related to the course.		5
	This game is an adequate teaching method for this course.		5
Perception of learning	I prefer learning with this game to learning through other ways (e.g. other teaching methods).		4
	The game contributed to my learning in this course		5
	The game allowed for efficient learning compared with other activities in the course.		5
	The game contributed to teach software testing.		5