**Understanding User Experience in Games: Evaluation between a Innovative Adaptive SmartController and a Traditional Controller**

Daniela Trevisan¹, Guilherme Gonçalves¹ and Érica Mourão¹

¹Fluminense Federal University, Computing Institute. Rua Passos da Pátria 156, Niterói, Rio de Janeiro, Brazil.

{daniela,galves,ericamourao}@ic.uff.br

**Abstract.** Emerging technologies offer exciting new ways of interacting with digital games to create fantastic play experiences. Evaluating entertainment technology is challenging because success isn’t defined in terms of productivity and performance, but in terms of enjoyment and gameplay experience. In this direction, the success criteria here has been evaluated in an exploratory user experience and usability study that used a within non-players design investigating a commercial joystick as input device compared to a novel adaptive touch-based controller. This new adaptive controller is able to dynamically change buttons size and position according to the user behavior. The evaluation methodology is based on the capture and analysis of objective and subjective user’s data. Results of this exploratory study indicate that while the general user experience and usability were similarly perceived in both controllers, the physiological measures indicate that the user's emotions were greater with the adaptive controller. The findings are subsequently discussed and implications of using the adaptive touch-based controller as input for digital games are discussed.

**Keywords:** User Experience based approaches; Evaluation Methods / Usability Evaluation; Human Factors and HCI; Adaptive Interfaces.

# 1. Introduction

When someone is playing games, two most important characteristics that defines the user perception about its experience are the quality and the fluence on gamepad [7] that are responsible for make the actions in-game. Due to this, the design of these devices are a very important factor in projecting an engaging gaming experience.

In the last few years, game studios and developers invested in the production of technologies to follow the trends, which are in a moment of rapid expansion and deep technological change [9]. Nintendo’s Wiimote[[1]](#footnote-0) and Microsoft’s Kinect[[2]](#footnote-1) enlightens the exploration of new ways of interaction through movements and gesture capture devices, and even smartphones can be well received by the desired target audience [5].

Digital games constitute an extremely varied set of applications with a rich range of experiences offered to players. This diversity makes difficult to devise a unique approach to their conceptualization and measurement. Terms such as fun, flow and gameplay are widely used to explain the user experience in game design [1]. However, there is an open discussion to include other relevant factors to the games. Emotion is often cited as a key element of the user experience [25]. Emotions in digital games act as a motivator for the cognitive decisions players make during gameplay and they drive user experience in digital games [15].In this direction to reach success in a digital game is related to providing a successful user experience with it. Part of such gaming experience success is regarding to the input device used to control the game interaction. Control is the player learning to manipulate the game. It is about the player learning how the objects in the game move, understanding the goals of the game and keeping the player occupied. It is also learning about the controllers, getting used tothe objects and angles in which the objects are displayed, and the ability of the player to memorise the relationship between controllers and the actions of the game.

While some work has been conducted regarding the development of input devices and how they affect user performance [5, 6, 18, 20] the effects of game controllers on user experience is yet to be explored in deeply.

Recently Torok et al [4, 7] propose a novel adaptive controller for digital game called SmartController[[3]](#footnote-2). SmartController with its adaptive interface consists of a mobile *app* available for Android e iOS. It is able to dynamically change buttons size and position according to the user’s behavior. User’s behavior is here interpreted as user’s touches on the screen.

Motivated by the demand from innovation in games controllers and the lack of techniques to assess the player experience with those devices,in this work the product success criteria is evaluated in an exploratory user experience and usability study that used a within non-players design investigating a commercial joystick as input device compared to a novel adaptive touch-based controller (e.g. SmartController).Our goal is to advance the theoretical understanding of user experience and controllers in digital games. In this work we are focusing on the measurement of user experience, usability and physiology in a game. However, one limitation of current psychophysiological studies is that they cannot precisely classify UX in digital games since many experiential phenomena in digital games lack standardized quantitative measurements [12]. Hence, another purpose of this study is to provide a correlation study between subjective user experience (estimated with AttrakDiff questionnaire) and objective physiological data (collected with EEG sensors).

# 2. How to evaluate UX in games

The current ISO definition on user experience focuses on a person’s perception and the responses resulting from the use or anticipated use of a product, system, or service. User experience includes all the users' emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors and accomplishments that occur before, during and after use [REF]. It is a vague definition and from a psychological perspective, these responses are actively generated in a psychological and emotional evaluation process, and it has to be decided which concepts can best represent the psychological aspects in games to allow to measure the characteristics of user experience.

Furthermore, user experience is a highlighted research topic inside the HCI community, which became present in many different fronts and themes, including digital games where the successful relationship between user and product is considered quite seriously. In literature exists many works whose goal is perform user experience evaluations in games [1]. In addition there are several ways to execute those evaluations, e.g. user tests by collecting objective and subjective data. In the following text we highlighted researches using questionnaires to capture the subjective perceived UX in games as well as those using psychophysiological data to derive the emotional state of player. Special attention is given to those researches dedicated in evaluating the players’ experiences with game controllers.

## 2.1 User Experience Questionnaires

Many works [12,15] have investigated different components of game experience using the game experience questionnaire (GEQ) [24]. The questionnaire was developed on the basis of focus group research [30] and following investigations among frequent players. The questionnaire consists of different modules: 1) Core module - concerns actual experiences during game play; 2) social presence module - concerns gaming with others; 3) post game module - concerns experiences once a player has stopped gaming. For instance in Nacke *et.al*. [15] participants played a fast-paced, immersive first-person shooter (FPS) game modification, in which sound (on/off) and music (on/off) were manipulated. As a result they found significant main effect of sound on all GEQ dimensions (immersion, tension, competence, flow, negative affect, positive affect, and challenge) concluding that subjective measures could advance the understanding of sonic UX in digital games.

Another widely used questionnaire is the Player Experience of Need Satisfaction (PENS)[ REFS games using PENS ]. The proprietary PENS-questionnaire, developed by Ryan and colleagues [37] investigates the intrinsic motivation of players [37]. This approach is based on self-determination theory (SDT) and focuses on the three basic human needs for either intrinsic or extrinsic motivation: Autonomy(volitional aspects of an activity), Competence (perception of challenge) and Relatedness (connection to others). Two additional factors - Presence (the sense that one is within the game world) and Intuitive controls – were included in the PENS measurement because they are seen as important aspects of gaming [37]. The PENS-questionnaire measures these needs and the additional factors on 7-point Likert-scales.

Recent research [14] investigated the factorial structure of both questionnaires PENS [37] and GEQ [24]. They found the factor structure of the PENS appears to be consistent and invariant across two different games, the GEQ reveals weaknesses in fulfilling these requirements.

Brown *et al.* [26] explored the experience, functionality and usability through standard and innovative gamepads, in order to evaluate these components. Subjective Mental Effort Questionnaire (SMEQ) [31] and Consumer Product Questionnaire (CPQ) [32] were respectively used in their work to measure effectiveness and satisfaction as subjective methods of usability evaluation. In other hand, Critical Incidents Technique (CIT) [33] was used to collect qualitative data, describing the user feelings and perceptions about its experience. CIT consists of a detailed analysis discourse technique.

In addition attractiveness has been used to measure UX in games [25,36]. Attractiveness in UX is described as a set of four dimensions: (1) Pragmatic Quality (PQ), (2) Hedonic Quality - Stimulation (HQS), (3) Hedonic Quality - Identity (HQI) and (4) Attractiveness (ATT) [11, 23]. The first one (PQ) represents traditional usability aspects, i.e. efficiency, effectiveness and learnability. It focuses on task related design aspects and indicates if the users reached their goals on a interaction. Following, the Hedonic Quality dimensions (HQS and HQI) describes quality aspects, which are not directly related to the tasks the user wants to accomplish with the software, for example originality and beauty. Both qualities are subjective aspects of a user interface. Thus, users may differ in their evaluation of these aspects. For instance in Lankees *et. al*. [25] they applied the AttrackDiff questionnaire to understand how emotional stimuli (facial expressions by Embodied Conversational Agents and emotion-eliciting situations) in interactive system affect the (more general) user experience. In another work [36]….

## 2.2 Psychophysiological Measures and its Correlation with Subjective User’s Data

The use of objective data collecting, like physiological measures (e.g. galvanic skin response, muscles contraction, respiratory and cardiovascular signals), are widely employed in the literature to evaluate the UX and the user engagement in digital games [12, 13, 15]. Moreover, the user emotional and mental state data can be retrieved from these measures, being useful in many recent researches in HCI community.

Some works have focused on studying the correlation between subjective users response and objective physiological data while assessing the player experience. For instance, in the work described by Mandryk et al [13] they collect galvanic skin response (GSR), electrocardiography (EKG), electromyography of the jaw (EMG), and respiration. Heart rate (HR) was computed from the EKG signal, while respiration amplitude (RespAmp) and respiration rate (RespRate) were computed from the raw respiration data. In their first experiment, they found many inconsistent correlations across participants. The main reason for the inconsistent results is likely the experimental manipulation that was chosen, however, there were also some methodological issues that contributed to irregular patterns of physiological activity. Primarily, the act of conducting the experiment produced different phases in the experiment (e.g. play, interview, rest) that created greater physiological responses than the experimental manipulations themselves. In addition, the experimental manipulation that was chosen (e.g participants played in four different conditions of difficulty: beginner, easy, medium, and difficult.) did not produce consistent subjective results across all participants. Without consistent subjective results, it is not possible to expect consistent physiological results. They observed the participants were responding more to the experimental situation than the experimental manipulations. Mainly methodological issues that had influenced those results were: high resting baseline, order and interview effects, subjects enjoyed playing in all conditions. In a second experiment they decided for maximizing the user experience and now the participants played in two conditions: against another co-located player, and against the computer. This time they could normalize and correlate the data showing that the amount by which participants increased their subjective ratings corresponded to the amount by which their mean physiological data increased.

Nacke *et.al*. [15]tried to correlate the GEQ dimensions with electrodermal activity (EDA) and facial muscle activity (EMG) to assess the player experience while music or sound is present or absent in a FPS game. They did not find any main effects of either sound or music on tonic psychophysiological data. Therefore, as in Mandryk *et. al* [13] they discussed whether physiological measures are indicated to assess such complex stimuli. There are many factors potentially affecting the interpretation and experience of perceptual stimuli: prior experience of the stimuli, cultural and societal experience and milieu, age, gender and present mood, for example.

# 3. Method

## 

## 3.1. Research Goals and Hypotheses

The goal of our research is to evaluate the user experience and usability in the use of a virtual adaptive control (SmartController) and traditional control in a digital game. Our approach is unique and unpublished in that it is the first, in our knowledge, to evaluate the user experience in the adaptive control versus the traditional control. The motivation of research goal is the demand from innovation in games controls and of how to evaluate it. We aim to study the user experience (e.g. attractiveness and emotions) and usability factors in the use of controls and the correlation between subjective users’ responses and physiological and performance data. In more detail, our researcher questions are:

**RQ1)** How the user experience (analyzed by users’ attractiveness and emotion) and usability are perceived in the use of SmartController and traditional joystick in the digital game?

Firstly we assume that a separate analysis of the game controllers could have an effect on the user experience dimensions and usability factors for the non-players users. Then the following hypotheses were formed:

**H1:** The user experience (e.g attractiveness and emotion criteria) with SmartController is greater than with the traditional one.

**H2:** Overall, the user preference is greater on SmartController rather than the traditional one.

**H3:** Overall, the game performance for non-players is similar in both controllers.

**RQ2)** Which dimensions of user experience (measured with Attrakdiff questionnaire) correlate with user emotions (measured with physiological EEG data) in the use of SmartController and traditional joystick?

Secondly, we are looking for potentials correlations between Attrakdiff data and EEG data. For instance, here we are interested in answering questions such as:

**H4**: User’s emotions are similar in both controllers during the experiment.

Lastly, other insights can be derived for further investigation of the interrelations between different indicators captured. They are:

1. Are the user’s emotions affecting the game performance?
2. Are the user’s emotions and user experience correlated?
3. Are the user preference and attractiveness dimensions correlated?

## 3.2 Evaluation Criteria

Our approach is based on the capture and analysis of subjective and objective user’s data. In order to understanding how the user experience is perceived through two game controllers usage we applied two methods: the AttrakDiff questionnaire [11, 27] used to analyze the attractiveness criteria and EEG data [21] used to estimate the user’s emotional state. Besides that, user preference is captured by applying the System Usability Scale (SUS) [28] questionnaire and the player performance is analyzed from game data log.

### 3.2.1 Attractiveness

Observing the nature of many questionnaires existent in literature (see the Section 2.How to evaluate UX in games) to evaluate user experience in games we chosen to use AttrakDiff based on two main reasons: (a) in most of the questionnaires are focused on measuring gameplay aspects and it is difficulty to separate quality aspects only related to the game controllers’ interaction. (b) AttrakDiff has been effectively used in various studies to investigate the perceived pragmatic and hedonic quality of many kinds of interactive and innovative products [23,25,29,36].

The model presented by Hassenzahl [11] assumes that pragmatic and hedonic quality are two independent quality factors of an interactive product. In addition, the model splits the hedonic quality into the two aspects stimulation and identity. Stimulation focuses on the human need for personal development, i.e. the need to improve personal skills and knowledge. A product can support this human need, for example, by providing new and stimulating functionalities. Identity focuses on the human need to identify itself on a product and the feeling of connection to it. At last, the attractiveness dimension fits to the global perception of how attractive a product is to the users [23].The AttrakDiff questionnaire [27] represents this attractiveness dimensions model. It contains twenty-eight items paired by words with a semantic differential scale, where each of these pairs means a questionnaire question [23]. The scales have seven scores and range the interval of -3 to 3. Also, the questionnaire questions are divided into seven items to each dimension.

### 3.2.2 User`s Preference and Performance

Usability and satisfaction or user preference do not exist in an absolute sense, which means they are strictly linked to their context [28]. The System Usability Scale (SUS) questionnaire was designed to evaluate the overall user satisfaction during a software system interaction [28]. The SUS score are calculated from all questions answers. It is desirable to obtain a mean score above 68, because a score below that indicates usability problems [28]. In addition, the SUS questionnaire was applied in our experiment to all participants after each controller test as a post-test evaluation method.

Looking at some research into controllers in general reveals that the performance is highly linked to the development process of these input devices [18, 20, 26]. Our focus in this context was to verify how the performance between both controllers is perceived. We used four metrics to represent the performance in-game: time spent in each stage, quantity of “lifes” at the end of the stage and the quantity of deaths in each stage. Moreover, we also collected the number of help requests from each participant.

### 3.2.3 Emotion Detection

Emotions in a psychophysiological context can be understood as connected physiological and psychological affective processes, which can be induced by perception, imagination, anticipation, or action triggers [34]. Perceptual emotions can be triggered by sensory information, such visual, acoustic, tactile, olfactory, or gustatory signals [35]. This distinction between emotional triggers is especially relevant when analyzing psychophysiological reactions together with game metrics.

The raw EEG signals were recorded with the Emotiv EPOC+ device [21] using the Xavier software platform, responsible for managing the recording sessions, processing and generating statistical reports from 10 participants. The device has 16 electrons in total, where 14 are used for data capture and 2 for reference and positioning. The EEG signals can be transmitted via usb or wireless connection to a receiver computer, where they are processed to generate reports with the mean value of each emotion (Engagement, Excitement, Interest, Relaxment, Stress and Focus) [16, 17].

Engagement and Excitement are emotions with mixed characteristics. Excitement is experienced as a consciousness or feeling of physiological excitement, i.e. animation, nervousness, anxiety. Engagement is experienced as alertness and conscious attention, directed to some stimuli, i.e. vigilance, concentration, stimulation, interest. This emotion is characterized by increased physiological excitation and EEG Beta waves along with EEG Alpha waves attenuation [22].

# 4. Experiment

## 4.1. Participants

Data were recorded from 10 volunteers with little or no experience in digital games. Participants were recruited in the University campus and no financial compensation was given. Their age ranged from 18 to 30 years. Six participants were male. None participants had previous experiences with the chosen game and most of them informed an experience above two years with smartphones. We also asked about the smartphone usage for gaming, where only 3 of 10 informed do not play games on smartphones. About the game console usage, 4 0f 10 informed they play games on a console at least twice a week, 2 of 10 informed a usage up to six days a week and others 4 informed they do not play on game consoles. Furthermore, we also collected their game gender preference resulting in (70%) Adventure, (50%) Strategy, (40%) Casuals, (40%) Simulation, (30%) Fight. Other genders took a percentage between 10-20%. The participants could choose more than one gender. We performed 3 pilot tests in order to improve the experiment.

## 4.2. Procedure

## 4.2.1. Experiment screenplay

The experiment proceed on the following steps. First, we applied the agreement term to the participants. This term requested permission to use the data obtained during the experiment, containing audios, images and video recordings, in order to produce technical and scientific papers always guaranteeing the anonymity. After that, we introduced the profile mapping questionnaire with the purpose to obtain personal data from participants. We fixed the Emotiv Epoc helmet as follows with the objective of capturing the physiological data during the test session. Before the game experiment starts, the participant was requested to close his eyes and concentrated his attention on the music playing during two minutes. Since physiological data has very large individualdifferences, this step serves as individual baseline capture. After that the participant received briefly instruction of how to interact with each controller. Finally, after theses steps the game experience started which took 10 minutes to each controller (as seen in Figure 1). At the end, the participants filled the SUS and AttrakDiff questionnaires for each controller experience. The alternance usage of controllers between participants was applied to avoid the learning effect of game mechanics.



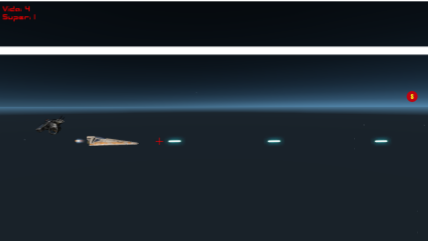
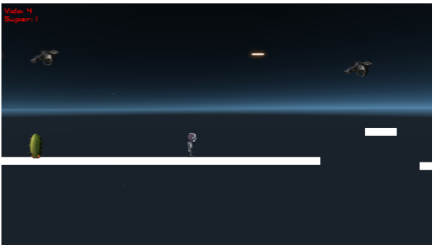
(a) (b)

Figure 1. Input devices used in the users sessions. (a) Participant playing the game with the SmartController and using the Emotiv headset. (b) Traditional Xbox controller.

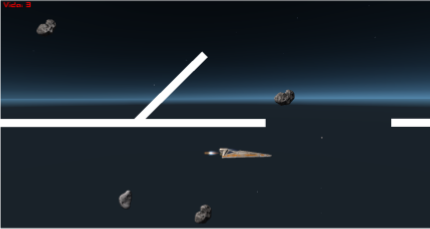
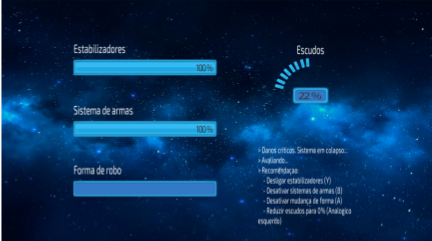
### 4.2.2. The Game

For the experiment section was selected a game with both controllers compatibility, called Guardian of Eternity, The objective is to escape the planet with the main character. It’s a simple game, with two stages. The player just has to reach the far right of the stage to finish it, avoiding getting hit by enemies or colliding with obstacles. Each hit decreases its health bar (or kills it instantly on stage 2). The player’s character starts with 3 lifes and have the possibility to gather more ‘life orbs’ during the stage. If the player dies, he will return from the last checkpoint. In stage 1, the player controls a robot and has to perform simple *platforming* tasks to defeat enemies and avoid hazards. The controller uses a simple layout with two directional arrows and a jump button, plus a transform button. Pressing the latter one changes the character to the spaceship form. Now the gameplay is a classic dual-stick shooter and the controller layout changes to two analog sticks (left one to move the ship and right one to shoot at any direction). Figure 3 shows the game stages screens, while Figure 4 show the respective adaptive controller layouts used to play [30].

After completing the stage 1, the main character is too damaged to keep going and it is necessary to disable several of its weapons and powers. In Figure 2 (c) we can see the minigame screen, where the player has to turn off the switches on the controller and turn down the power using the dial. This is an example of the kind of direct and natural interaction that this controller provides (as shown on Figure 3 (c)), mapping a real-world action to an interface that is way more similar to real devices than mere buttons on a gamepad. The player, now stuck in the spaceship form, proceeds to enter a maze, where he has to avoid colliding with walls or asteroids. The weakened guardian has troubles to fight the effects of gravity. The player must constantly press "up" to avoid falling to the ground. The controller (as seen in Figure 3) has buttons to controls the vertical and horizontal movements separated in the opposing sides of the interface, an attempt to provide a more ergonomic alternative to a regular d-pad [30].



1. Robot at Stage 1 (b) Spaceship at Stage 1



(c) Minigame (d) Spaceship at Stage 2

Figure 2. Game screens for both stages. Stage 1 has two gameplay modes: platforming as robot (a) and dual-stick shooter with the spaceship (b). After finishing this level, the player enters a minigame (c) that precedes the second stage (d).



1. Robot at Stage 1 (b) Spaceship at Stage 1



(c) Minigame (d) Spaceship at Stage 2

Figure 3: The layout displayed by the SmartController during the different stages of the game.

## 4.3. Equipaments

The experiments were made in a controlled environment in a computing institute room at Fluminense Federal University. Were used the Emotiv Epoc+ (Model 1.1) for physiological recordings, a Xbox controller device, a smartphone (Nexus 5) used as gamepad and two notebook computers, one Core i7 3th generation CPU, 8 GB of RAM and a Nvidia GeForce 745M GPU for running the game and other Core i7 6th generation CPU, 16 GB of RAM and a Nvidia GeForce 940M GPU for receive the Epoc+ data stream. The operating system was Windows 10.

# 5. Results

For ease of presentation and hence understanding, we present the experiment results in the following subsections. Besides, we applied a Wilcoxon signed ­rank test, a non­parametric statistical hypothesis test with significance level of 0.05 and the two-tailed hypothesis defined, returning the p-value, which indicates if the difference between the results achieved for both groups is significant. The decision of using a non­parametric test was motivated by the unknown distribution of the test results, since a parametric test requires a previous knowledge about the data distribution, which must be normal. For our groups' sizes, a p-value smaller than 0.05 represents a significant difference. Furthermore, for correlation results we applied a Pearson linear correlation measure expecting to find some correlations between our variables.

## 5.1. Attractiveness

For the hypotheses H1 (The user experience (e.g attractiveness and emotion criteria) with SmartController is greater than with the traditional one.), the Figure 4a shows the obtained results from Attrakdiff in its four dimensions, with each the respectively means and standard deviations. The Pragmatic Quality dimension (PQ) is perceived as greater in traditional controller (M = 3.89, SD = 0.72) than the adaptive one (M = 3.99, SD = 1.21). Although, the Wilcoxon test for the hypothesis H1 returned a non statistically significant result, since p-value = 0.661. The Hedonic quality - identity (HQI) is also greater in traditional controller (M = 4.49, SD = 1.47) than the adaptive one (M = 4.40, SD = 1.03). Although, the Wilcoxon test for the hypothesis H1 returned a non statistically significant result, since p-value = 0,556. The Hedonic quality - stimulation (HQS) is also greater in traditional controller (M = 4.91, SD = 1.72) than the adaptive one (M = 3.79, SD = 0.81). Although, the Wilcoxon test for the hypothesis H1 returned a non statistically significant result, since p-value = 0,989. The Attractiveness dimension (ATT) as well as the others, is greater in traditional controller (M = 3.93, SD = 0.71) than the adaptive one (M = 3.86; SD = 0.91). Although, the Wilcoxon test for the hypothesis H1 returned a non statistically significant result, since p-value = 0,743. As seen in Figure 4a, all results found for hypothesis H1 returned a non statistically significant result. The Figure 4b show the mean in scale Likert -3 to 3.

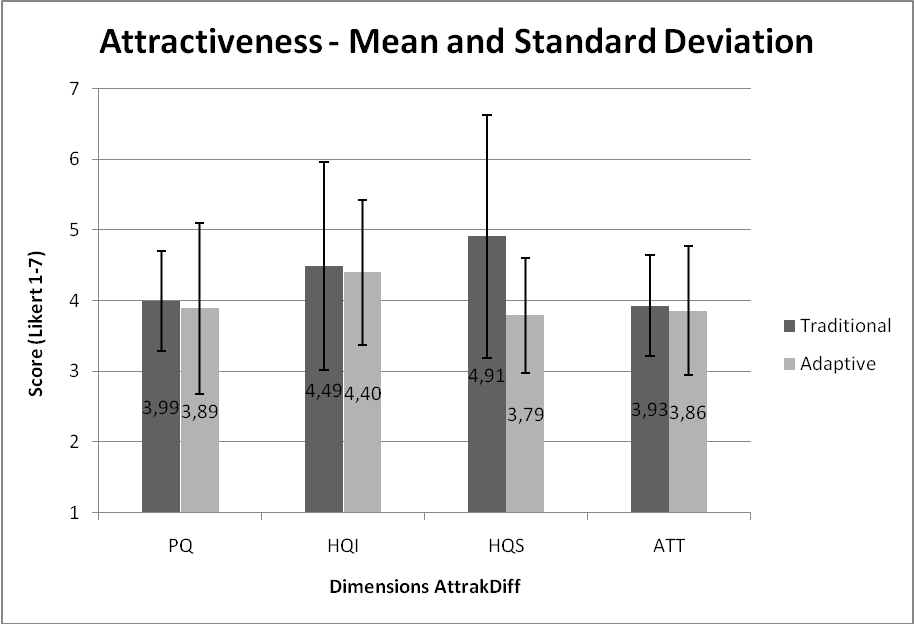


Figure 4a:AttrakDiff Questionnaire Results – Mean and Standard Deviation

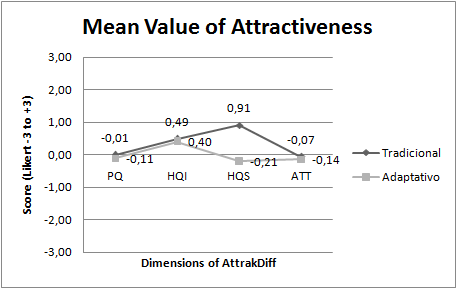


Figure 4b:AttrakDiff Questionnaire Results – Mean in Likert -3 to 3

## 5.2. User´s Preference

For the hypothesis H2 (Overall, the user preference is greater on SmartController rather than the traditional one.), the results show that the UX during the experiment is perceived as lower in adaptive controller than traditional controller. Comparing the mean score from SUS in adaptive controller (64.8) and the traditional controller (74.0) with 5% of margin of error, we perceived that the score in adaptive controller is below the desirable mean (68.0) . When we paired the participants SUS scores, the Wilcoxon results show a non significant p-value of 0.853, rejecting the hypothesis H2.

## 5.3. Performance

For our hypothesis H3 (Overall, the game performance for non-players is similar in both controllers.) we have the results obtained according to performance metrics (Stage duration in seconds, quantity of remaining player’s lifes at the end of stage and number of player’s deaths). For the metric stage duration in seconds (stage 1) was observed a lesser value in traditional controller (M = 145,4) than the adaptive one (M = 158,7). In Stage 2 the traditional controller is greater (M = 268,3) than the adaptive (M = 243,9). Although, the Wilcoxon test obtained on this metric for both stages, respectively, resulted on a p-value of 0.375 and 0.444. This results shows a significant difference between the results for this analysis. Figure 5 illustrates the duration averages (in seconds) of each game stage and their respective standard deviations.

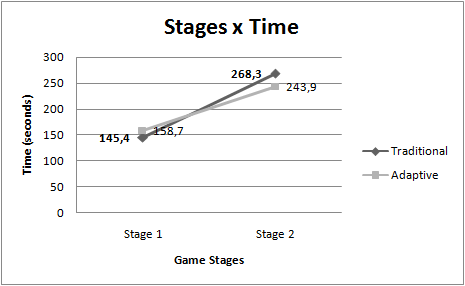


Figure 5: Mean Time (seconds) for each game stages.

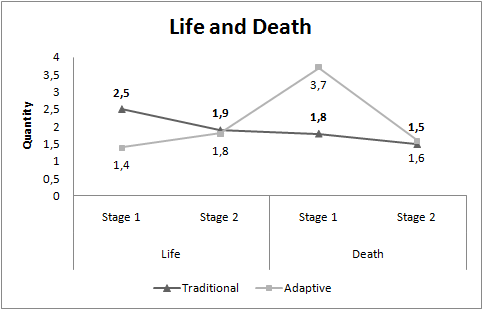


Figure 6: Mean Time (seconds) for each game stages.

Quantity of remaining life at the end of stage 1 is perceived as greater in traditional controller (M = 2.5) than the adaptive (M = 1.4) and at stage 2 also is greater in traditional controller (M = 1.9) than the adaptive (M = 1.8). However the number of player deaths is greater in adaptive controller in stage one (M = 3.7) than traditional (M = 1.8) and in stage second (M = 1.6) than (M = 1.5) respectively. Regarding the metric (quantity of remaining life at the end of each stage), we found a significant difference only for the first one, obtaining a p-value of 0.020, while in the second one, we had a non-significant p-value of 0.821. For the metric (number of player deaths), no significant difference was evidenced in both stages, since the p-values where 0.188 and 0.930. In this scenario, the only statistically significant result indicates a performance improvement in stage 1 from the game using the adaptive control. The Figure 6 show us the results.

## 5.4. Emotion analysis

For the hypothesis H4 (User’s emotions are similar in both controllers during the experiment.), we compared the emotions (Engagement, excitation, interest, relaxation, stress and focus) results between both controllers. Our findings show that all emotions are greater in adaptive controller. The engagement in adaptive controller scored higher levels (M = 57.5, SD = 3.98) than the traditional controller (M = 57.4, SD = 4.30). The excitement also got greater levels in adaptive controller (M = 20.3, SD = 7.10) than the traditional one (M = 18.1, SD = 6.62). The interest, in the same way was greater in adaptive controller (M = 56.1, SD = 3.73) than the traditional one (M = 55.7, SD = 3.06). The relaxation as well as the others was greater in adaptive controller (M = 31.2, SD = 3.22) than traditional one (M = 31.1, SD = 2.51). The stress was perceived as greater in adaptive controller (M = 48.0, SD = 12.96) than traditional one (M = 46.7, SD = 11.79). Finally, the focus was also greater in adaptive controller (M = 34.8 SD = 9.76) than traditional controller (M= 32.6, SD = 8.26). Overall emotions are greater in adaptive controller. We verify if exists significant difference. The Figure 7 shows the means and standard deviations from each emotions recorded from Emotiv epoc + device. After results from Wilcoxon show a significant p-value (< 0.05) only in excitement (0.0141) and focus (0.0245) emotions.

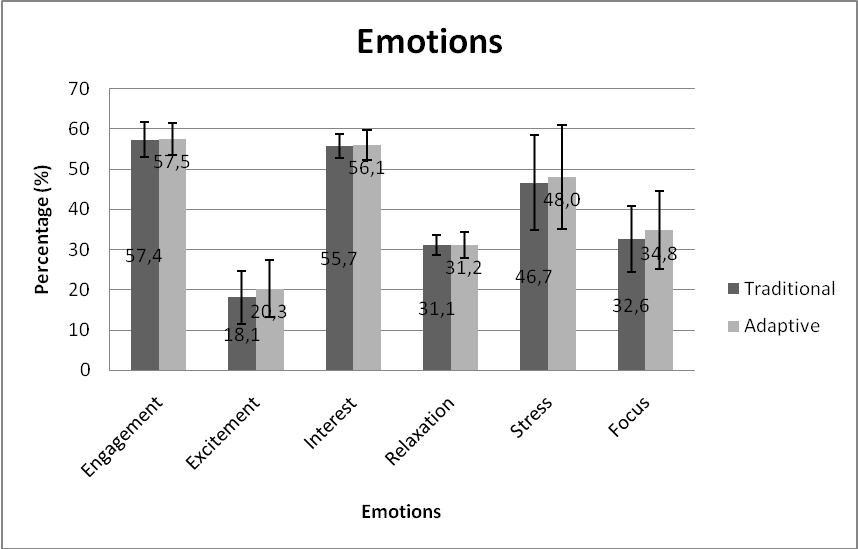


Figure 7: Emotions - Mean and Standard Deviation.

We observed, according to table 1, that the excitement and focus emotions are those that obtained p-value below 0.05, which indicates that there is significant difference. For the other emotions in this study, i.e. engagement, interest, relaxation and stress, no significant difference is evidenced, which means that these emotions are similar in both controllers.

|  |  |
| --- | --- |
| **Emotions** | **p-value** |
| Engagement | 0,8539 |
| Excitement | **0,0141** |
| Interest | 0,5245 |
| Relaxation | 0,8539 |
| Stress | 0,3997 |
| Focus | **0,0245** |

Table 1: Wilcoxon p-values for each emotion.

## 5.5. Insights about interrelations of emotions and performance

Regarding the two variables, we investigate whether the user’s emotions are affecting the game performance. For that, we tested the correlation between performance and emotion using the Pearson linear correlation, which a strong correlation is defined as a p-value lower than 0.05. A significative difference in number of player’s deaths metric was found in stage 1. The Pearson value from excitement and interest emotions and this performance metric shows a moderate correlation (as seen in Table 2).

|  |  |  |  |
| --- | --- | --- | --- |
| **Emotion versus Number of player’s deaths** | | | |
| Emotions | **Controller** | **P-value < 0,05** | **Pearson Correlation**  **0,5 <= r < 0,8** |
| Excitement | Traditional | 0,0065 | 0,73 |
| Adaptive | 0,0161 | 0,78 |
| Interest | Traditional | 0,0242 | 0,69 |
| Adaptive | 0,0670 | 0,59 |

Table 2. Emotions versus Number of player’s deaths in the Game (Only statistically significant (p < 0.05) results are reported.)

For stage 2, the emotions "excitement" and "interest" have significant differences regarding the metrics "stage duration in seconds" in the use of both controllers (see table 3). Therefore, the Pearson coefficient evidenced a moderate positive value for "interest" and a strong positive value for "excitement", which indicates a correlation for each emotion.

|  |  |  |  |
| --- | --- | --- | --- |
| **Emotion versus Stage duration in seconds** | | | |
| **Emotions** | **Controller** | **P-value < 0,05** | **Pearson Correlation**  **0,5 <= r < 0,8 moderate 0,8 < = r < 1 stronger** |
| Excitement | Traditional | 0,0016 | 0,85 |
| Adaptive | 0,0044 | 0,81 |
| Interest | Traditional | 0,0466 | 0,64 |
| Adaptive | 0,0322 | 0,67 |

Table 3. Emotions versus Stage duration in seconds

Brain waves identified in Brain Activity Map are: Delta, Theta, Alfa and Beta for all sensor location on Emotiv Epoc headset. Delta waves (0.5-4 Hz) indicate, when active, deep sleep, rest and, conversely, agitation when suppressed. Theta (4-8Hz) are waves that indicate state of deep, dream meditation. Alpha waves (8-15Hz) indicate states of relaxed alertness, rest and meditation. And Beta waves (15-30HZ) indicate wakefulness, alertness, mental engagement, conscious information processing [19]. Some difference in the EEG might be explained by a mixed of emotions. An increase in stress might also fell some kind of excitation [20, 18].

The Figure 8 can be interpreted as follows: the red color in Theta wave is linked to vigil with forced attention state (concentration) to relaxation before the game. Alpha is linked to wakefulness with relaxation. Beta associated with vigil with state of attention.

The Figures 9 and 10 can be interpreted as: the hotter colors in Alpha and Beta indicates an increase in frequency in these waves.

The results show that the Alpha wave had an increased frequency in both controls at the time of the player's death but the activity was the same in adaptive and traditional controller.

However in the Beta wave increase was greater in adaptive control than in traditional. Beta indicates that there was an increase in activity associated with intense mental concentration. The emotion excitement can be explained by the increase in Beta frequency. As the red coloration is lower in the traditional control for the Beta wave, we understand that the excitation was lower.

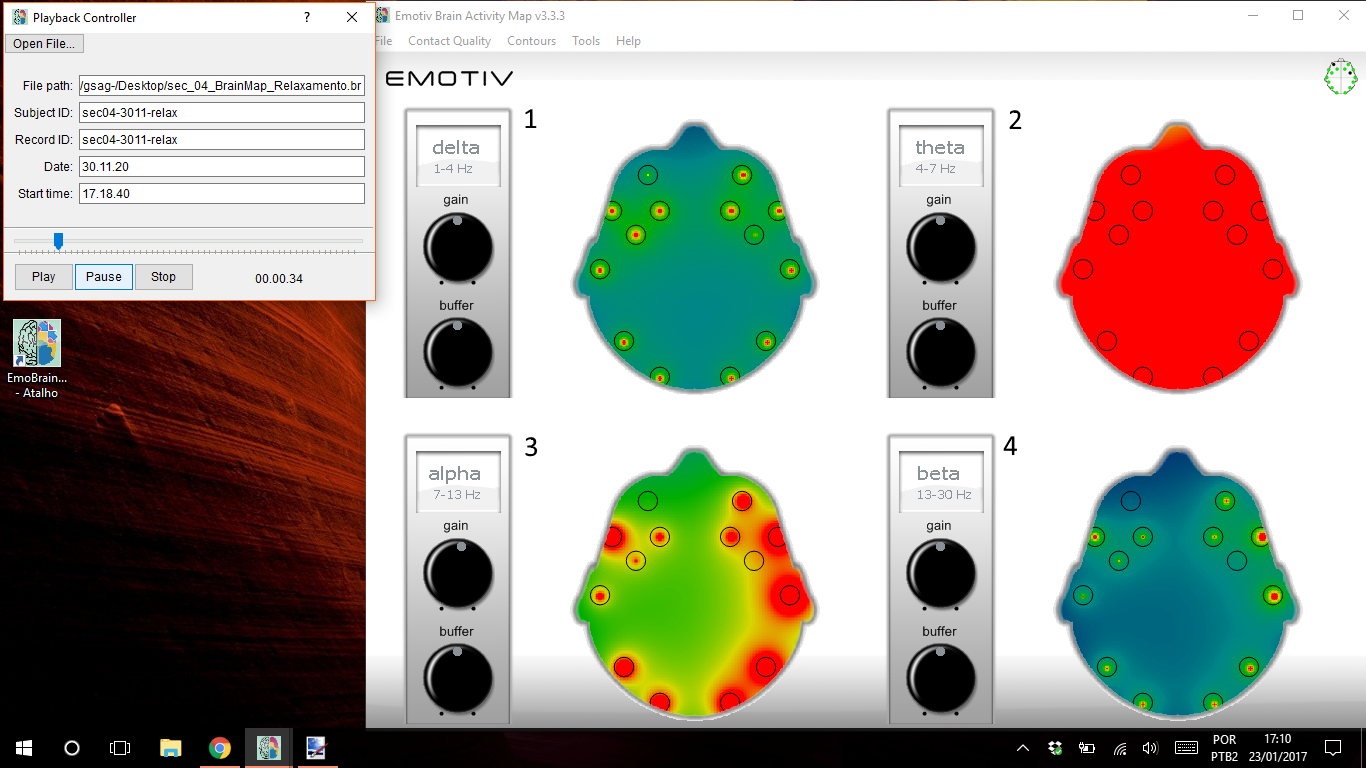


Figure 8 – Relaxation before the game (Neural baseline)



Figure 9 - Adaptive – Neural activity in the moment of player’s death.



Figure 10 - Traditional - Neural activity few seconds before player’s death.

## 5.6. Insights about interrelation between emotions and user experience

For investigate whether the user experience are correlated with the user’s emotions, we take the results from AttrakDiff dimensions and the emotions. The correlation between the Pragmatic Quality dimension and the Interest emotion was strong (p-value = 0.0066) with value 0.7894 of moderate correlation. There were no significant differences between other emotions that indicated correlation. With this, we can conclude that there is no correlation between the the hedonic quality (HQI) and the interest emotion.

## 5.7. Insights about interrelation between attractiveness and user preference

Further investigation of the interrelations between different indicators from attractiveness and user preference were computed using Pearson correlation coefficient. As a reference for usability and user experience attractiveness, respectively, the analysis included the items from the SUS questionnaire and the Attrakdiff questionnaire dimensions.

The SUS rating of the controllers showed a non significant correlation at p < 0.05 with all attractiveness dimensions. For the adaptive controller we obtained the following values (PQ -0.445, p = 0.197; HQI 0.182, p = 0.614; HQS -0.445 p = 0.197; ATT 0.449, p = 0.193) and for the traditional one we obtained the following values (PQ 0.283, p = 0.428; HQI 0.384, p = 0.273; HQS -0.254, p = 0.478; ATT 0.023, p = 0.949). In general, the results suggest that the attractiveness and user preference factors are weakly interrelated. It is important to investigate in more detail how the attractiveness influences the perception of the user preference, or if the two concepts are independent.

# 6. Discussion

The study reported here is the first experiment to compare an innovative controller and a traditional one. At current stage we have to deal with the following limitations: (1) the few amount of performance variables from the game and controllers; (2) the emotion and performance analysis could be complemented with other physiological measures i.e Electromyography (EMG). We believe that future works with more performance variables could bring some news insights. As well as, new physiological measures helped to try to better understand the results.

The results from Attrakdiff questionnaire in its four dimensions is realized as greater in traditional controller than adaptive controller. However the user experience, in the context of attractiveness, did not show a significative difference between the adaptive controller and the traditional one in our hypothesis (H1). Which means that the user experience in both controllers are perceived as similar.

The results of the hypothesis (H2) indicate that the user preference is greater in traditional controller than adaptive controller but has no significant difference between both controls. This shows that the user’s preference is also similar in the adaptive controller and the traditional.

The user experience and the user preference may have been affected by the participants' profile: for the inexperienced players, both controls are a new experience. We believe that greater game duration would bring more significant data.

We observed in the performance metric (stage duration in seconds) in hypothesis (H3) that the participants concluded each stage with different durations. The Wilcoxon test showed significant differences between controllers in stages durations. The first stage took more time with adaptive controller than traditional and the second stage took less time with adaptive controller than traditional controller. This leads us to think that in stage one the participant took more time to adapt to adaptive control and in stage two the user had already adapted to the touch.

The duration of second stage was greater than the first one for both adaptive and traditional controllers. This can be explained by the presence of many obstacles in the second stage of the game. To do not die the player needed to deflect of the obstacles using only directional buttons up and down.

Regarding the performance metrics (quantity of remaining player’s lifes at the end of stage) and (number of player’s deaths) in hypothesis (H3) we observed the results. The quantity of remaining player’s lifes at the end of each stage was greater in traditional control than adaptive control. The significant different was perceived only in stage one. Maybe the player adapting to the touch of adaptive control in stage on. With this, more life would be left in the use of traditional control. The number of player’s deaths was greater with the use of adaptive control than the use of traditional control in both stages. However, no significative difference was evidenced. In both controls, adaptive and traditional, the player died more times in stage one than second stage. We believe that the player paid more attention to the obstacles of the second stage game to die less.

In this scenario presented of the hypothesis (H3), the only statistically significant result indicates a performance improvement from stage one to stage two in the game using the adaptive control and we concluded that the performance is not similar in controls.

The work reported here show results that user’s emotions are not similar in both controllers during the experiment. The emotions of experiment (Engagement, Excitement, Interest, Relaxation, Stress, and Focus) are greater in adaptive controller than traditional controller but wilcoxon test show significant difference only in excitement and focus. We believe that the excitement was greater in adaptive control by being a different, new control in a smartphone handset. Already the emotion focus was greater on adaptive control because it did not have a tactile, physical button and overall, you would not need to look for control during the game. This way the player would remain focused on the game screen. Looking into the all emotions of the study we concluded to hypothesis (H4) that some user’s emotion are similar and other are not similar in both controllers.

The results show that the correlation analysis between emotions and performance exists. First a significative difference was found to emotions (excitement and interest). Pearson correlation value shows moderate correlation in emotions versus number of player's death in the stage one of the game. Also the emotions "excitement" and "interest" have significant differences regarding the metrics "stage duration in seconds" in the use of both controllers in second stage. Therefore, the Pearson coefficient evidenced a moderate positive value for "interest" and a strong positive value for "excitement". This results of first insight show that emotions excitement and interest affect the performance number of player's death and stage duration in seconds.

Results from AttrakDiff dimensions and the emotions show the moderate correlation between the Pragmatic Quality dimension and the “Interest” emotion. Others dimensions and emotions not have correlation for second insight. This may indicate that the interest raises when the quality of an application raises and the degree of success that users achieve the goals.

Additional analysis between SUS rating and AttrakDiff dimensions in third insight shows in general that user preference and attractiveness do not correlate. The profile of the participant with little or no experience and the style of the game could have influenced the results. An in-depth study could bring more significant results.

There are many factors that affect the user experience. The experimental study investigated gameplay experience using subjective (AttrakDiff and SUS questionnaire) and objective measures (psychophysiological - EEG and evaluate performance). Clearly, these results present need to find a more different forms of evaluation to more insight.

Unfortunately, we did not found more significative aspects between the adaptive and traditional controllers that evidenced others conclusive answers to our research questions. We hope with our study that the contribution made may inspire other researchers to explore this field of science.

# 6. Conclusion and future works

In summary, we have found out that evaluate the user experience is not an easy task. Through this first comparative study between a traditional controller and the adaptive one, we collected data from 10 participants while they were playing with both controllers. Participants played the game while measures such EEG (subjective responses) was taken and questionnaire responses (objective responses) were collected. A methodological result of the research presented in this paper indicates that while the general user experience and usability were similarly perceived in both controllers, the physiological measures indicate that the user's emotions were greater with the SmartController. In addition, the correlation of psychophysiological activity and subjective responses will be an analysis strategy to concentrate on in future studies. We also see a lot of potential in cross-correlation of subjective and objective measures in terms of attentional activation (i.e. brain waves electroencephalogram - EEG ). In the future works we hope to investigate the relationship between physiological and questionnaire responses in more detail. This study shows that it was focused on non experienced users in games and thus it might be hypothesized that the results are only valid for this group. It remains for future research to indicate that measurements can describe user experiences for a broader demographic population. Certainly we will taken more results in the future.

# Acknowledgements

We would like to express our gratitude to Adriel Araújo, Bruno Olímpio and José Santos for helping in this study. We also would like to thank all our participants.

# 8. References

1. Bernhaupt. R.  2015. Game User Experience Evaluation. Human Computer Interaction Series, Springer.
2. Hotho, S., McGregor, N. 2013. Changing the Rules of the Game: Economic, Management and Emerging Issues in the Computer Games Industry. Springer.
3. YoYo Games: GameMaker. http://www.yoyogames.com/gamemaker. Accessed 23/11/2016.
4. Torok, L., Pelegrino, M., Lessa, J., Trevisan, D.G., Vasconcelos, C.N., Clua, E. and Montenegro, A. 2015. Evaluating and customizing user interaction in an adaptive game controller. Design, User Experience, and Usability: Interactive Experience Design, DUXU 2015. LNCS, vol 9188. Springer, pp. 315-326.
5. Malfatti, S. M., Dos Santos, F.F. & Dos Santos, S.R. 2011, Using mobile phones to control desktop multiplayer games, Proceedings - 2010 Brazilian Symposium on Games and Digital Entertainment, SBGames 2010, pp. 230.
6. Vajk, T., Coulton, P., Bamford, W., Edwards, R., 2007. Using a Mobile Phone as a “Wii-like” Controller for Playing Games on a Large Public Display, International Journal of Computer Games Technology 2008*,*e539078.
7. Torok, L., Pelegrino, M., Trevisan, D.G., Clua, E. & Montenegro, A. 2015, A mobile game controller adapted to the gameplay and user’s behavior using machine learning. Entertainment Computing - ICEC 2015, Springer, pp. 3-16.
8. Pelegrino, M., Torok, L., Trevisan, D. & Clua, E. 2014, Creating and Designing Customized and Dynamic Game Interfaces Using Smartphones and Touchscreen, Brazilian Symposium on Games and Digital Entertainment, SBGAMES, pp. 133.
9. VentureBeat, 2016. PwC: Game industry to grow nearly 5% annually through 2020. Available at: http://venturebeat.com/2016/06/08/the-u-s-and-global-game-industries-will-grow-a-healthy-amount-by-2020-pwc-forecasts/. Accessed: 20/01/2017.
10. Steam Community: GestureWorks Gameplay. <https://steamcommunity.com/app/296610>. Accessed 20/01/2017.
11. Hassenzahl, M. 2005, Hedonic, emotional, and experiential perspectives on product quality. In Encyclopedia of Human Computer Interaction, pp. 266-272.
12. Nacke, L., E., Grimshaw, M., N., Lindley, C., A., 2010, More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. Interacting with Computers, Springer, pp. 336-343.
13. Mandryk, R. L., Inkpen, K. M., & Calvert, T. W. 2006. Using psychophysiological techniques to measure user experience with entertainment technologies. In Behaviour and Information Technology. vol. 25 no. 2, pp. 141-158.
14. Brühlmann, F. and Schmid, GM. 2015. How to Measure the Game Experience?: Analysis of the Factor Structure of Two Questionnaires. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15). ACM, New York, NY, USA, 1181-1186.
15. Nacke, L., E. and Lindley, C. 2009. Affective Ludology, Flow and Immersion in a First- Person Shooter: Measurement of Player Experience. The Journal of the Canadian Game Studies Association. Vol 3, No 5, pp. 1-21.
16. Ramirez, R. and Vamvakousis, Z. 2012, Detecting Emotion from EEG Signals Using the Emotive Epoc Device. In Brain Informatics, Springer. pp. 175-184.
17. Cernea, D., Kerren, A., and Ebert, A. 2011. Detecting Insight and Emotion in Visualization Applications with a Commercial EEG Headset. In Conference on Evaluations of Graphics and Visualization - Efficiency, Usefulness, Accessibility, Usability, pp.53-60.
18. Zhang, Z. Microsoft Kinect Sensor and Its Effect. 2012. IEEE Multimedia, vol. 19, no. 2, pp. 4-10.
19. Emotiv: Epoc Brain Activity Map. <https://www.emotiv.com/product/epoc-brain-activity-map/>. Access 18/11/2016.
20. Natapov, D., Castellucci, Steven J. and MacKenzie, I. Scott. 2009. ISO 9241-9 evaluation of video game controllers. In Proceedings of Graphics Interface 2009 (GI '09). Canadian Information Processing Society, Toronto, Canada, pp.223-230.
21. Emotiv. <https://www.emotiv.com/> Access 18/11/2016.
22. Emotiv: User Manual for Emotiv Xavier Control Panel Epoc. <https://www.emotiv.com/>. Access 18/11/2016.
23. Schrepp, M., Held, T. and Laugwitz, B. 2006. The influence of hedonic quality on the attractiveness of user interfaces of business management software. In Interacting with Computers, 18, 5, pp. 1055-1069.
24. IJsselsteijn, W., van den Hoogen, W., Klimmt, C., de Kort, Y., Lindley, C., Mathiak, K., Poels, K., Ravaja, N., Turpeinen, M., and Vorderer, P. 2008. Measuring the experience of digital game enjoyment. In Proc. of Measuring Behavior, Maastricht, Netherlands, pp.88-89.
25. Lankes, M., Bernhaupt, R., Tscheligi, M. 2015. Evaluating User Experience Factors using Experiments: Expressive Artificial Faces Embedded in Contexts. In Game User Experience Evaluation p:113-133. Springer.
26. Brown, M. A., Kehoe, A., Kirakowski, J. and Pitt, I. 2015. Beyond the Gamepad: HCI and Game Controller Design and Evaluation. In Game User Experience Evaluation, pp. 263-285, Springer.
27. AttrakDiff. <http://attrakdiff.de/index-en.html>. Access 08/12/2016.
28. Brooke, J. 1996. SUS-A quick and dirty usability scale. Usability evaluation in industry, v. 189, n. 194, p. 4-7.
29. Bernhaupt, R. Desnos, A., Pirker,M., Schwaiger. D. TV Interaction Beyond the Button Press - Exploring the Implications of Gesture, Pressure and Breath as Interaction Mechanisms for a TV User Interface. IFIP TC13 Conference on Human-Computer Interaction 2015 (INTERACT). p:1-8. Springer.
30. Poels, K., de Kort, Y.A.W., & IJsselsteijn, W.A. (2007). It is always great fun! Exploring dimensions of digital game experience using focus group methodology. Proceedings of the 2007 Conference on Future Play (Toronto, Canada, 14-18 November 2007), 83-89.
31. Arnold, A. 1999 Mental effort and evaluation of user interfaces: A questionnaire approach. In: Proceedings of 8th International Conference on Human-Computer Interact, ACM, New York, pp. 1003–1007.
32. McNamara, N. and Kirakowski, J. 2011. Measuring user-satisfaction with electronic consumer products: The Consumer Products Questionnaire. In: International Journal of Human-Computer Studies, 69, 6, pp. 375-386.
33. Flanagan, J. C. 1954. The Critical Incident Technique. Psychological Bulletin. pp. 1-33.
34. Nacke, L. E. 2013. An Introduction to Physiological Player Metrics for Evaluating Games. In Game Analytics: Maximizing the Value of Player Data. pp 585-619.
35. Bradley, M. M., & Lang, P. J. 2007. Emotion and motivation. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), Handbook of psychophysiology (3rd ed., pp. 581–607). New York: Cambridge University Press.
36. Christou, G. 2012. Exploring player perceptions that contribute to the appeal of World of Warcraft. In Proceedings of the 4th International Conference on Fun and Games (FnG '12). ACM, New York, NY, USA, 105-108.
37. Ryan, R., Rigby, C., and Przybylski, A. 2006. The motivational pull of video games: A self-determination theory approach. Motivation andEmotion 30, 4, pp. 344-360.

1. Wiimote: http://wii.com/ [↑](#footnote-ref-0)
2. Kinect: https://developer.microsoft.com/en-us/windows/kinect [↑](#footnote-ref-1)
3. SmartController: http://smartcontrollerapp.com/ [↑](#footnote-ref-2)