A Practical Evaluation of the Influence of Input Devices on Playability

Lucas Machado(✉) and João Luiz Bernardes Jr.

School of Arts, Sciences and Humanities (EACH),
Universidade de São Paulo, São Paulo, Brazil
{lucas2.machado,jlbernardes}@usp.br

Abstract. Innovations being achieved with interactive devices (screens, sensors etc.) allow the development of new forms of interaction for many applications. Videogames played with these devices are completely changing how we use them and taking advantage of intuitive interfaces. Based on that, we ask "What aspects of playability are affected using different input devices for a certain gaming task and how is gaming performance affected?". Our contribution is to present a practical evaluation of four different input devices (Mouse, Gamepad, Kinect and Touchscreen) used to interact with the same game, Fruit Ninja, with our data analysis indicating that changing input device brings significant differences in certain aspects of player experience for this game, such as sensation, challenge and control, while for others there was very little difference since this particular game rarely provides intense experiences for those aspects.

Keywords: Emotions in HCI · Entertainment systems · Evaluation methods and techniques · User experience · Input devices

1 Introduction

Videogames nowadays are relevant both as an economic activity and as a subject of scientific research. Economic relevance is due to, for instance, the large growth of this sector, particularly when compared to the average growth of other areas and of national economies as a whole [1]. Scientific relevance is shown in game research in several areas and applications, such as Human Computer Interaction studies or gamefication techniques being applied in education, among many other examples. Research based on videogames has a ludic nature which keeps user attention for longer periods and provides experiences such as exploration, challenge, competition, and collaboration [2], that are different from the experiences expected from traditional productivity software.

While playability relates to user experience in videogames, there is not a *de facto* definition for it. However, as the traditional usability evaluation heuristics are used with user interfaces in productivity software, several authors have tried to propose frameworks to evaluate Player Experience in videogames, considering the particular properties of this medium [3]. Thus, in the same way that usability

M. Kurosu (Ed.): HCI 2016, Part III, LNCS 9733, pp. 385–395, 2016.
DOI: 10.1007/978-3-319-39513-5_36

in software aims to improve users productivity with a set of heuristics related to ease-of-use and simplification of interfaces and tasks [4], proposals of playability heuristics aim to improve game experience with characteristics like challenge, fantasy, humor, exploration etc. [2].

Human Computer Interaction happens through input and output devices and, as Bonarini et al. [5] notes, the interaction with devices turns out to be more important than perceptual realism of the game in how strong a relationship can be established between the game and the user. Also as Alvarez et al. [6] points out, over the last decades the devices we use to interact with games have evolved considerably, and these changes shaped our usage and, thus, our experience of interaction and play. These studies show opportunities for research on the area of input devices and playability, like the work of Lee and Chung [7], that analyses psychological experiences of gamers, in which the results demonstrated that the controller was significant in the perceived interactivity, spatial involvement, dynamic immersion, and realistic immersion. However, there are still few studies and a gap in literature relating input devices to specific game experiences, especially if compared to usability studies of traditional productivity software, and a question can be elaborated of top of it: "What aspects of playability are affected by using different input devices for a certain gaming task, and how is the gaming performance affected?".

Our objective is to address this question by selecting criteria and metrics of playability evaluation in the literature that relate to interactions and input methods and also to the subjective aspects of player experience, preparing and carrying out observational experiments with a particular videogame, and establishing what are the most appropriate metrics for evaluation of interaction of users with games, under certain devices and tasks, and then conducting an experiment and analyzing the data to investigate whether different devices could produce different game experiences, to establish a framework to facilitate the understanding of the relationship between playability and input devices.

An experiment was then conducted, in which voluntary users were asked to play a game alternating between four different input devices: Mouse, Gamepad, Touchscreen and Microsoft Kinect. After the experiment, the users answered a questionnaire that collected data about their experiences. One limitation is that the chosen game (*Fruit Ninja*) does not have the capacity to strongly generate all the possible experiences (e.g. eroticism or fellowship), and a more extensive experiment is required to analyze them under longer gaming sessions (probably in an immersible Role Playing Game that could emulate all these experiences, or using different games for different experiences). This limitation does, however, provide an opportunity to test the quality of experimental results, since we expect the experiences not evoked by the game to be ranked uniformly low by users and, thus, to show little difference between input devices. We are also purposefully focusing on these experiences separately from questions related to ergonomics or the physical strain of prolonged use, which is why the experiment is composed of only a few short, one-minute play sessions (which is the default duration for one of Fruit Ninja's modes).

As a contribution, our results could help in game development processes, for example by choosing the best interaction devices and techniques based on the desired experiences. The process of evaluation itself can also be reproduced for other games or devices and even during development and can be considered another contribution. Further studies could also help in considering player experiences during controller design.

After a statistical analysis of the data collected on the experiment, differences between user experiences in different devices were found. As expected, several emotions showed very little difference between devices or any intensity at all during this experiment. Sensation, Fantasy, Challenge, Expression and Control, in that order were the aspects which showed the most significant differences between different devices. Kinect was the preferred input device, followed by touch.

2 Related Work

There is not a consensus yet about a most accepted set of heuristics for playability evaluation, but there are some promising studies trying to define it. The set of Heuristic Evaluation for Playability (HEP) by Desurvire et al. [8] is focused on *expert evaluation* during the development phase of a game, as well as on aspects related to gameplay, game story mechanics and usability. Since these heuristics are more direct and objective, and they are not strongly linked to player feelings, it was not adequate to our purposes. The set of heuristics described by Koeel et al. [9] is based on the previous cited HEP and on *expert evaluation* as a test framework and it is also strongly related to game mechanics.

Other research exists relating user experiences and interactions to games and input and output devices. Lee and Chung [7] studied gamer experiences according to display and controller, but focusing more on physical or materially perceived experiences comparing 2D to 3D environments. Bonarini et al. [5] also studies wearability and comfort of devices and their physiological effects on users, and while this could relate to player experience with different devices, it is still distant from players emotions and feelings, that could be mapped do playability heuristics.

Chu et al. [10] proposed a playability matrix based on an analysis of methodologies for evaluating player experience in game play. It was used in this work in conjunction with the previous cited literature research to define our framework of evaluation, based on the playful experiences (PLEX) framework. As described by Lucero et al. [2], the PLEX framework summarized heuristics of playability evaluation of several other theoretical works on pleasurable experiences, game experiences, emotions, elements of play and reasons why people play. Since it deals with experiences as feelings and emotions themselves, it is the best fitted set of heuristics to abstract from the experiences we want to identify on players with different devices.

Thus, since there are no studies specifically treating and comparing the relation of different input devices and user experiences and emotions in games, we try to fill this gap by using PLEX heuristics to evaluate a set of devices under determined tasks.

3 Evaluation of Input Devices

Since different input devices provide different user interactions, different devices could be needed to keep the usability of a system high depending on the user interface and the tasks to be executed. We ask then, for a specific set of tasks and interfaces, how different input devices affect the playability of a game.

Therefore, our objective in the experiment is to propose and establish playability metrics and criteria to evaluate certain input devices (mouse, gamepad, touchscreen and Kinect) for certain tasks in a game.

3.1 Methodology

An experiment was conducted with 13 voluntary users in a controlled environment, in which they were asked to play the game *Fruit Ninja* in arcade mode (for the period of one minute, only once after a training game), alternating between four different input devices: Mouse, Gamepad, Touchscreen and Microsoft Kinect. After the experiment, users anonymously answered a questionnaire ranking their device preference, previous experiences with the devices, and the intensity of their experiences in several aspects with each device. Not all of the emotions and experiences described in PLEX are expected to be present in a game like *Fruit Ninja*, but we chose to use the questionnaire without modification and let the collected data show the truth of this. Therefore, we chose to evaluate experiences of Captivation, Challenge, Competition, Completion, Control, Cruelty, Discovery, Eroticism, Exploration, Expression, Fantasy, Fellowship, Humor, Nurture, Relaxation, Sensation, Simulation, Submission, Subversion, Suffering, Sympathy, and Thrill. The experiment sessions were previously scheduled to prevent conflicts and interruptions and a period of one hour was reserved for each session. Despite the relatively low number of participants, experimental results showed good statistical significance.

Risks. During this experiment, participants were exposed to minor physical discomfort risks, such as aspects of posture, reading distance, illumination and noise, or dirtiness or contamination through devices they touched. Another possible risk is embarrassment by skill confrontation, since the experiment protocol could be interpreted to ask some participants to perform above their skill levels. To mitigate all of these risks, devices were cleaned after each use, adaptations to specific physical conditions of the participants were available, and participants were not encouraged to pay attention to their in-game scores and asked to execute the task in the way they felt most comfortable. They were also told they could ask questions and interrupt or leave the experiment at any time.

Participants. We had thirteen participants recruited in our university. They were mainly students or professors. The average age was 24.5 years (ranging from 17 to 38, with a standard deviation of 6.86), and 7.7 % of them identified themselves as women (92.3 % identified as men).

Table 1. Participants previous experiences on devices

Experience	None	Little	Average	A lot
With Fruit Ninja game	15.4 %	23.1 %	53.8 %	7.7 %
With touch screen	7.7 %	0 %	23.1 %	69.2 %
With mouse	15.4 %	7.7 %	15.4 %	61.5 %
With gamepad	15.4 %	23.1 %	46.2 %	15.4 %
With Kinect	38.5 %	30.8 %	15.4 %	15.4 %

As we can see in the Table 1, the devices with which participants had more previous experience were touchscreen, mouse, gamepad, and Kinect, in decreasing order. More than half of the participants knew the game well and only 15.4 % of them had never played it. Touchscreens were also reported as a more common device than the mouse, perhaps because of their current extensive use in smartphones. Participant experience with gamepad was more varied, probably because videogames have been in the market for decades and, despite not being as ubiquitous as smartphones in Brazil, they are still quite popular. The Kinect device, on the other hand, was unknown or little known for most participants, probably due to the fact that it is a rather new device and available only in a specific platform (the Xbox console).

3.2 Instruments

For each of the forms of input being evaluated and to answer the questionnaire, several devices were used. In the Kinect evaluation, an Xbox 360 console with a Kinect attached to a video projector were used. The setup was that the projected screen size was about 32 in. on its diagonal, and the user was about 4 meters from it. The image was never obstructed by any of the devices. The projector (or a large screen) was necessary because the users have to be at a certain minimum distance from the kinect so it can capture their image, and playing the game in a smaller monitor from this distance would be impractical. For the touch screen evaluation, a third generation iPad was used. In both mouse and gamepad evaluation, a 13-inch MacBook Air running Windows 10 in a virtual machine environment was used, with a common mouse and an Xbox One controller wired via USB. The participants played on the computer screen.

The questionnaire was also answered on the computer. We implemented a simple web service with a wizard interface so the participant could go through the questions and assign a value for the perceived intensity for each of the PLEX categories and for every device. Visual aids and descriptions about the experiences were used to help identifying and understanding them: Captivation (Forgetting one's surroundings), Challenge (Testing abilities in a demanding task), Competition (Contest with oneself or an opponent), Completion (Finishing a major task, closure), Control (Dominating, controlling, regulating), Cruelty (Causing mental or physical pain), Discovery (Finding something new or unknown), Eroticism

(A sexually arousing experience), Exploration (Investigating an object or situation), Expression (Manifesting oneself creatively), Fantasy (An imagined experience), Fellowship (Friendship, communality or intimacy), Humor (Fun, joy, amusement, jokes, gags), Nurture (Taking care of oneself or others), Relaxation (Relief from bodily or mental work), Sensation (Exciting by stimulating senses), Simulation (An imitation of everyday life), Submission (Being part of a larger structure), Subversion (Breaking social rules and norms), Suffering (Experience of loss, frustration, anger), Sympathy (Sharing emotional feelings), and Thrill (Excitement derived from risk, danger). Because all users were native speakers of Brazilian Portuguese, this was the language actually used in the questionnaire.

To adapt gamepad input signals to the game, we used Xpadder software. It maps and simulates controller input as keyboard and mouse. For our tests, we mapped the left analog stick as mouse movement, and the "A" button as a left click. That way, the user should simulate mouse movements by pressing "A" while moving the stick to simulate the cutting blade movement.

For the game installations, we used the official versions of *Fruit Ninja* for iOS and Kinect. For the tests on the computer (with mouse and gamepad), we used an unofficial version that was ported to PC but has the same dynamics and graphics, except for the pomegranate bonus in the end that was missing. Fruit Ninja was chosen because it is possible to run it under different platforms for different input devices and because the game itself is simple to understand and has an arcade game mode that lasts for one minute, standardizing the sessions between players and devices.

3.3 Experimental Procedure

After agreeing to an Informed Consent document, in which participants could also read about aspects of the experiment pertaining to themselves, such as its objectives and risks, they were asked to interact with the game in arcade mode for one minute and only once per device, as the test task. Participants could remove their consent or leave the test at any moment without having to give any explanations. Initially we randomly chose the order of the first three devices between mouse, touchscreen and gamepad. We set Kinect always as the last one since it requires greater physical effort and this could influence the evaluation of other devices.

Before each device evaluation, participants could try the device in a training session in the same game mode (arcade) for one minute. It was an optional choice and they often decided to skip the training if they already had previous experience with the game or with the device.

At the end of the experiment, participants were asked to answer the questionnaire. Participants had to fill in their age, gender identity, and their previous experience with the game and with each of the devices used in the experiment. For these questions, a scale of one to four was used, with one being *no experience* and four *plenty of experience*. This was followed by a series of 22 questions (one for each experience according to the PLEX framework [2] we chose). Participants were asked to evaluate the intensity of their experiences in a scale of zero to ten for each device, for every experience, with zero denoting no intensity and ten full intensity. Finally, participants were asked to sort the devices by how much they enjoyed each to execute the tasks in the game. An optional field allowed participants to write any relevant comments they wanted to express. After each session, devices were cleaned and games were restarted to prepare for a next session.

4 Results

Data collected from the questionnaire answers were extracted and the Analysis of Variance (ANOVA) method was applied to analyze the scores participants provided to the aspects of their experience for each of the devices.

The results show differences between user experiences in different devices. As expected, several emotions (half of the set of experiences) showed very little difference between devices or any intensity at all during this experiment. Sensation, Fantasy, Challenge, Expression, Control, Humor, Thrill, Suffering, Simulation, Captivation and Exploration, in order from most to least difference, were the aspects which showed significant differences between different devices.

Table 2. F-values and P-values of variables with significant difference

	F-value	P-value
Sensation	10.27027027	0.0000245473656
Fantasy	8.528064993	0.0001208619902
Challenge	8.521854305	0.0001215729081
Expression	8.296525097	0.0001505365208
Control	7.230474732	0.0004240317507
Humor	6.399884426	0.0009783774656
Thrill	5.867	0.001696429971
Suffering	5.838978015	0.001746782684
Simulation	4.710676447	0.005819916856
Captivation	3.269876003	0.02907170146
Exploration	3.161100196	0.03292371113

With our data and an alpha of 0.05, an F-critical of 2.798 was calculated in the ANOVA procedure. Figure 1 shows a comparison of the averages of each device for the experiences that had a F-value greater than the F-critical, according to the values of Table 2.

There was also no substantial correlation between preferred devices and previous experience with them. As Fig. 2 shows, Kinect was the preferred input device, followed by touch. Participants disliked using the gamepad a lot, since it was considerably harder to play the game with it, and a few of them reported

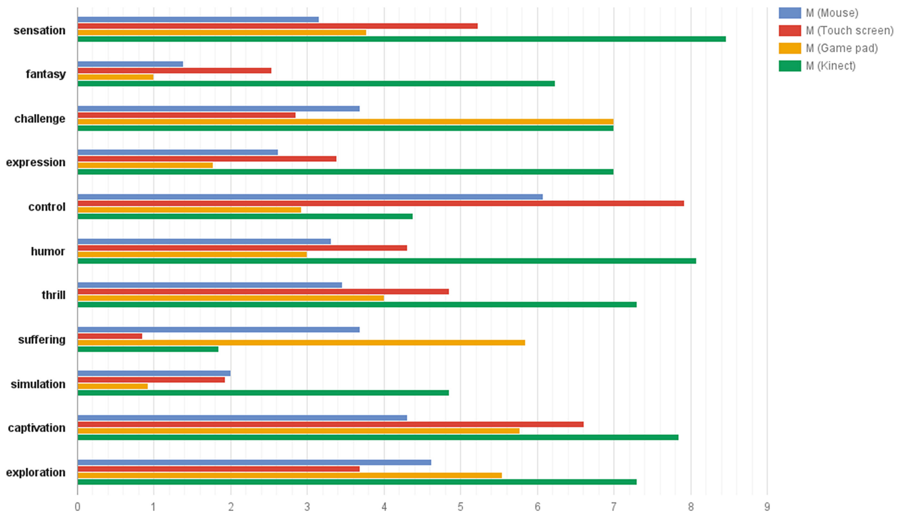

Fig. 1. Comparison of averages between devices for the most significant differences in experiences

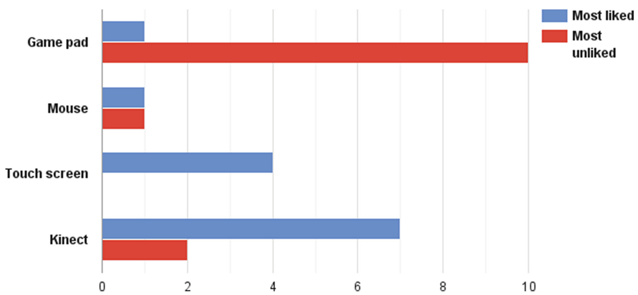

Fig. 2. What devices participants liked and disliked the most

disliking Kinect because they felt it gave them less control over the game due to lack of physical contact with the device.

The raw data obtained in the experiment is available at https://drive.google.com/file/d/0B7274IzRmrS8OS13dVBWbk42NVU.

5 Discussion

We reached our initial objective of evaluating players experiences through different devices. It was clear that almost every experience was more intense when using the Kinect device, except for Control and Suffering. We believe this is related to the use of the whole body (particularly the arms in this case) to control the game, which even allowed the use of both hands, together or separately, in an intuitive way and occasionally led to humorous situations. Other factors may have affected the experience with Kinect. Some of them are common to most videogame experiences with this device, such as necessitating a larger screen and standing some distance from it. Other differences, related to our experimental setup, appear to have had little or no impact. While game versions for Xbox, iPad and PC were different, gameplay, graphics, sound and other features were very similar in all of them except for the method of input. We also investigated whether the lack of experience with this device could be responsible for its high scores, due to some sort of novelty effect, but there was little correlation between these scores and experience with Kinect (and users with average or high experience also scored high on the relevant experience aspects when using it). We believe the aspect of control was low for the Kinect because interactions using it occur without physical contact, so users have no tactile feedback and rely solely on visual and aural cues about how the game reacts to their input. The low intensity of Suffering on Kinect might be caused by the higher intensity of other positive aspects which help prevent emotions of anger and frustration. Touchscreen provided the least Suffering experience, since users were manipulating the game directly and could feel that difficulties they had originated from themselves and not from the device interaction (incidentally, the game was initially developed with this device in mind).

It is also worth noting the high intensity of Challenge and Suffering with the gamepad, since it is a device best suited to directional controls and hard to use in the chosen game. However, The intensity of Captivation and Exploration were also high for this device, suggesting that participants were interested by the difficulties and in learning how to use it for this game.

Touch screen demonstrated the highest intensity of control when comparing to the other studied devices. This is probably because it is a direct manipulation device in which users could directly *feel* their fingertips *cutting* the fruits on the screen. Mouse followed touchscreen closely in Control experience, perhaps because its use is common among all the participants and it feels as natural and easy to control the cursor as with direct manipulation. Despite Kinect also looking realistic and natural with its movements recognition, the game control is indirect since the user has to command a shadow of himself on the screen, therefore not achieving the same level of Control experience. However, Fantasy intensity on Kinect was noticeably high among other devices, being explained by the realism of doing body movements that looked like a ninja (the game's avowed theme).

Overall we can observe that the Kinect device could improve the intensity of player experiences and feelings for games of this type, especially fantasy and humor, but that it reduces the feeling of control (which might be desirable or not depending on the desired game experience). The Gamepad device offers difficulty and challenge since it is not best suited for the chosen game. Mouse tends to follow touchscreen, usually with slightly less intense emotions, and players feel good control on touchscreens since its direct manipulations feels natural.

As this work was conducted with only one game, Fruit Ninja, some experiences such as eroticism and fellowship could not be easily aroused on the participants.

An experiment with a longer gaming session using a game which has the capacity of stimulating all the experiences, such as an immersible and social RPG, or several smaller games, each stimulating different experiences, would be necessary to have a wider understanding of their relations with input devices.

Other input devices were also considered to be for inclusion in this study, but left out due to time and resource limitations. Future work could try to reproduce the experiments with other games and devices such as Leap Motion, Intel RealSense, graphics tablets, game paddles, or even voice control.

Future work in the design of input devices and in game designing could also benefit from our contribution and the contribution of further experimentation along this line in what concerns which specific experiences the designer aims to provide. The better understanding of the relationship between playability and input devices allows the translation of usability concepts to the field of videogame studies and design, but relating to player experiences and feelings and ways to analyze them.

We expect that more studies along these lines will be made since interaction device technology is improving quickly and also becoming more pervasive. In a world of different devices that collect data and commands from users for games, it is valuable to understand how users feel and react to these devices.

References

1. Makuch, E.: Video game industry grows four times faster than US economy, ESA says (2014). http://www.gamespot.com/articles/video-game-industry-grows-four-times-faster-than-u/1100-6423536/
2. Lucero, A., Holopainen, J., Ollila, E., Suomela, R., Karapanos, E.: The playful experiences (plex) framework as a guide for expert evaluation. In: Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces, DPPI 2013, p. 221 (2013). doi:10.1145/2513506.2513530, http://dl.acm.org/citation.cfm?d=2513506.2513530
3. Paavilainen, J.: Critical review on video game evaluation heuristics. In: Proceedings of the International Academic Conference on the Future of Game Design and Technology, pp. 56–65 (2010). doi:10.1145/1920778.1920787, http://dl.acm.org/citation.cfm?id=1920787
4. Turner-Bowker, D.M., Saris-Baglama, R.N., Smith, K.J., DeRosa, M.A., Paulsen, C.A., Hogue, S.J.: Heuristic evaluation of user interfaces. Telemed. J. e-Health Official J. Am. Telemed. Assoc. **17**, 40–45 (1990). doi:10.1089/tmj.2010.0114. CHI 2009. http://online.liebertpub.com//abs/10.1089/tmj.2010.0114
5. Bonarini, A., Costa, F., Garbarino, M., Matteucci, M., Romero, M., Tognetti, S.: Affective videogames: the problem of wearability and comfort. In: Jacko, J.A. (ed.) Human-Computer Interaction, Part IV, HCII 2011. LNCS, vol. 6764, pp. 649–658. Springer, Heidelberg (2011). doi:10.1007/978-3-642-21619-0-77
6. Alvarez, J., Haudegond, S., Havrez, C., Kolski, C., Lebrun, Y., Lepreux, S., Libessart, A.: From screens to devices and tangible objects: a framework applied to serious games characterization. In: Kurosu, M. (ed.) HCI 2014, Part III. LNCS, vol. 8512, pp. 559–570. Springer, Heidelberg (2014). doi:10.1007/978-3-319-07227-2-53

7. Lee, H., Chung, D.: Influence of gaming display and controller on perceived characteristics, perceived interactivity, presence, and discomfort. In: Kurosu, M. (ed.) HCII/HCI 2013, Part II. LNCS, vol. 8005, pp. 258–265. Springer, Heidelberg (2013). doi:10.1007/978-3-642-39262-7-29
8. Desurvire, H., Caplan, M., Toth, A.F.: Using heuristics to evaluate the playability of games. In: CHI 2004 Extended Abstracts on Human Factors in Computing Systems, pp. 1509–1512 (2004). doi:10.1145/985921.986102, http://portal.acm.org/citation.cfm?d=985921.986102
9. Koeffel, C., Hochleitner, W., Leitner, J., Haller, M., Geven, A., Tscheligi, M.: Using heuristics to evaluate the overall user experience of video games and advanced interaction games. In: Bernhaupt, R. (ed.) Evaluating User Experience in Games: Concepts and Methods, pp. 233–256. Springer, London (2010). doi:10.1007/978-1-84882-963-3. http://link.springer.com/chapter/10.1007/978-1-84882-963-3_13
10. Chu, K., Wong, C.Y., Khong, C.W.: Methodologies for evaluating player experience in game play. In: Stephanidis, C. (ed.) Posters, Part I, HCII 2011. CCIS, vol. 173, pp. 118–122. Springer, Heidelberg (2011). doi:10.1007/978-3-642-22098-2