ASSESSING THE IMPACT OF DISPLAY SIZE, GAME TYPE, AND USAGE CONTEXT ON MOBILE GAMING QOE

Justus Beyer, Viktor Miruchna, Sebastian Möller

Quality and Usability Lab, Telekom Innovation Laboratories, TU Berlin

ABSTRACT

In this paper we present results from the Quality of Experience (QoE) evaluation of two commercially available games on four different smartphones and tablets with screen sizes between 3.27" and 10.1". The tests were conducted in a neutral lab and a simulated metro environment. Results show a considerable impact of display size on overall quality as well as four out of seven Player Experience dimensions. No significant impact of the usage context on gaming QoE was observed. Conclusions are drawn with respect to how experiments addressing gaming QoE should be set up to obtain generalizable results.

Index Terms— gaming, Quality of Experience (QoE), evaluation methods, display size, context

1. MOTIVATION AND INTRODUCTION

Computer games are becoming more and more popular. This refers not only to games running on traditional playing devices, but also for smartphones and tablets. According to surveys between a third and half of the time spent with smartphones is dedicated to gaming [1]. As these devices have not been dedicatedly developed for gaming activities and might not permit an optimal experience due to their limited input and output capabilities, this might come as a surprise, but points to the fact that the quality of gaming experience is not yet well understood. In contrast to pure media consumption, the user of a computer game actively interacts with the medium, which leads to different outcomes. Other than with task-oriented services, a gamer pursues his activity because it is rewarding in itself and quick termination is not desired, causing concepts such as effectiveness and efficiency to partially lose their positive value for overall quality. In turn, there are other quality aspects, which need to be considered, such as positive and negative emotions resulting from playing, or the flow experienced by the user. A taxonomy of quality aspects and related metrics is provided in [2]. In order to quantify such quality aspects in a valid and reliable way, experiments with human test participants are necessary, who rate perceived quality after each interaction on dedicated questionnaires (other methods to quantify QoE in gaming situations, such as physiological signals, are still under study). In this paper, we report on experiences gained during an evaluation of two smartphone games on four devices with different screen sizes in two simulated usage contexts. Whereas the quality judgments for the games are not of scientific interest, our analysis focuses on experimental factors affecting the obtained quality judgments. In Section 2, we describe the experimental set-up, followed by the test procedure in Section 3. Section 4 analyzes the results, which are then discussed with respect to the implications for future gaming QoE evaluations in Section 5.

2. EXPERIMENTAL SETUP

For our study in August 2013 we chose four popular screen sizes between 3.27" and 10.1". To minimize effects of differing hedonic device quality, we selected devices from one brand: Samsung Galaxy Young (3.27"), Galaxy S4 (5"), Galaxy Tab 3 7.0 WiFi (7"), and Galaxy Tab 10.1 (10.1"). Although their build quality and case materials are comparable, different display technologies are used and processing power differs. Yet, all were well capable to run the tested games without limitations.

As the usage context (physical and/or social) is expected to be a confounding influence on mobile gaming QoE, it has to be simulated as well. The first setting was a laboratory room following ITU Rec. [5] and [6] with participants sitting on an office chair next to a desk. The "metro" environment simulated a driving train with reduced lighting and train noises. The participant's space was limited using two grey partition screens very close to the sides of the player. In an effort to imitate the effects of a moving train, participants sat on a unsteady one-legged bar chair.

We chose two games based on their visual and input/control complexity: "Flipper Spiel Pinball" is a simple game representing a classic flipper with the player shooting a ball across a board using two levers, earning points by hitting targets. The more complex "Striker Soccer Euro 2012" is a soccer game where the player controls the actions of one team in a real-time soccer match, trying to score more goals than the other automatically controlled team.

3. TEST PROCEDURE

The study was conducted using a within-subject design with 26 participants (17m, 9f; 22y-48y, avg. 25.5y) having prior experience in mobile gaming. After being instructed about the purpose of the experiment and filling in an introductory questionnaire, examining demographic information and prior experience with games and interaction with smartphones and tablets, the participants had to play a total of 12 game scenarios of three minutes each in random order. After each test session, a three-part questionnaire had to be answered, containing the 46-item core part of the Game Experience Questionnaire (GEQ) [3], one question for overall quality, and 4 further questions examining the suitability of the game for the present display. These questions had to be rated on an ACR scale labeled according to [4].

The GEQ is designed to assess seven dimensions of Player Experience (Sensory and Imaginative Immersion, Tension, Competence, Flow, Negative Affect, Positive Affect, and Challenge). Of the 12 tested conditions, 8 were situated in the neutral environment (both games on each device) and 4 conditions took place in the simulated metro (both games, only the biggest and smallest device).

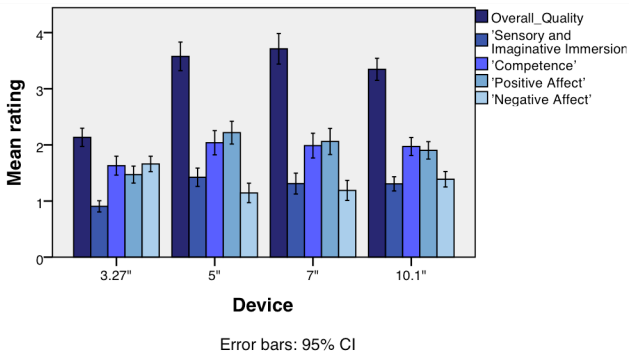

Fig. 1. Mean quality ratings of the tested display sizes

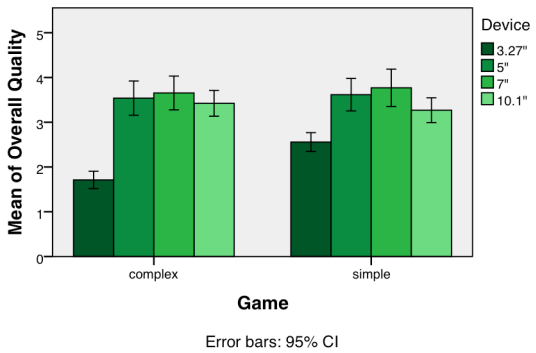

Fig. 2. Mean overall quality ratings for the tested games

Dimension	Sig.	F(3, 300)	η^2
Immersion	p < .01	11.41	0.10
Competence	p < .01	4.58	0.04
Positive Affect	p < .01	10.33	0.09
Negative Affect	p < .01	6.48	0.06

Table 1. Significant effects of the factor display size

Dimension	Sig.	F(1, 300)	η^2
Overall quality	p < .05	4.78	0.02
Competence	p < .01	33.44	0.10
Tension	p < .01	43.40	0.13
Challenge	p < .01	80.00	0.21

Table 2: Significant effects of the factor game

4. RESULTS

We analyzed the collected GEQ data and the overall quality ratings from 312 sessions using a MANOVA with the independent variables game, setting, and device and the dependent variables overall quality, sensory and imaginative immersion, competence, flow, tension, challenge, positive affect, and detail quality (suitability for display). The analysis showed that the overall quality MOS is significantly affected by display size ($F(3,300) = 38.87$, $p < .01$, $\eta^2 = .319$): Ratings using the smallest tested display size were significantly lower (Scheffé post hoc test) than using the other displays. Among these bigger screens no significant differences were found (see Fig. 1). Significant influences of the display size factor were also observed for the quality dimensions shown in Table 1. While these effects exist for both games, they are more pronounced for the complex game (see Fig. 2). Significant effects of the game factor are shown in Table 2. The environment factor showed no significant influence on any of the tested dimensions. However, one participant remarked that he felt more comfortable in the metro situation being hidden from the experimenter by the partition screens.

5. DISCUSSION AND FUTURE WORK

In this paper, we have presented the results of a study on the impact of the factors game type, display size, and context on perceived playing experience. Our results confirm that the display size has a strong influence on the perceived quality of a gaming session. Although the screen sizes used in the experiments were not equally spaced on a continuum, there seems to be no linear link of quality with size. Instead, it seems that an acceptance threshold is reached as soon as the display has reached a certain size (in our case 5"), and then quality and its sub-dimensions do not further increase significantly. We could show that smaller devices lead to lower playing experience ratings while gaming sessions with larger devices received higher marks. Considering the low ratings for Competence on the smallest device combined with the insignificance of the device's influence on the Challenge dimension, it seems that the increased difficulty of playing on a small touch screen is not perceived as a challenge but as annoyance, causing the observed higher Negative Affect scores. As initially assumed, small devices are better suited for playing the simple than the complex game. Although the games influenced ratings, the magnitude of the their impact on the overall quality was lower than expected. It is possible that the participants focused just on the display sizes. In our study the games' difficulty remained the same for all participants, potentially making them overly easy for some participants and too difficult for others. As the equilibrium of demanded skill and a player's abilities is a prerequisite for flow experience, games are needed that can be easily adapted towards a participant's capabilities. While lack of influence of the simulated "metro" environment might mean that no context effect exists, we remain doubtful of this: The "metro" simulation may have been insufficient in that it did not take the social context into account to an adequate degree. Although the experimenter never interfered with the participants' playing, he was visible and his observation perceivable for the player in the neutral environment, whereas he was hidden in the "metro" setting by the partition screens. Thus, further studies are needed to clarify the influence of the context on mobile gaming. A study of the social context as a factor may also be fruitful as the strong influence of social interaction on playing experience in non-computer games is evident.

7. REFERENCES

[1] S. Khalaf, Mobile Use Grows 115% in 2013, Propelled by Messaging Apps, http://bit.ly/khalaf, Jan. 13, 2014, accessed Feb. 25, 2014.
[2] S. Möller, S. Schmidt, J. Beyer, "An Overview of Concepts and Evaluation Methods for Computer Gaming QoE", QoMEX 2013, AT-Klagenfurt, 2013.
[3] W. Ijsselsteijn, Y. De Kort, K. Poels, F. Bellotti, and A. Jurgelionis, "Characterising and Measuring User Experiences in Digital Games", ACE 2007.
[4] ITU-T Rec. P.800, Geneva, 1996.
[5] ITU-T Rec. P.910, Geneva, 2008.
[6] ITU-T Rec. P.911, Geneva, 1998.
Co-funded by the German BMBF, funding code 01IS12056.