BOREDOM, IMMERSION, FLOW – A PILOT STUDY INVESTIGATING PLAYER EXPERIENCE

Lennart Nacke
Blekinge Institute of Technology, Game and Media Arts Laboratory
Box 214, SE-37424 Karlshamn, Sweden

Craig Lindley
Blekinge Institute of Technology, Game and Media Arts Laboratory
Box 214, SE-37424 Karlshamn, Sweden

ABSTRACT

Designing and evaluating gameplay experience comes to life after measures for player experience have been found. This paper describes a pilot study measuring game experience with a set of game stimuli especially designed for different player experiences. Gameplay experience is measured using self-report questionnaires after each play session. Results of the questionnaires are then separately compared to design intentions and player evaluations. Our experiment shows that gameplay experience can be assessed with a high reliability for certain gameplay features.

KEYWORDS

Gameplay experience, game design, self-report measures, quantitative study, player research.

1. INTRODUCTION

Playing digital games has become the most popular leisure activity in recent years. This demands interesting content and an immersive experience. Designing for this is difficult, as there are currently no reliable measurements of game experience. Thus, it is hard to base game design on empirical facts. The approach to game design presented here is based on measuring the player experience. So, we must analyze in detail what it is exactly that drives players to play digital games. This leads us to present a study on the experience evaluated by players of a first-person shooter level.

A variety of measures for play-related experiences in games have been proposed by other researchers (IJsselsteijn et al., 2007, Griffin, 2005) with focus in behavioral psychology. Many studies on games have assigned and studied the experience of gameplay as a part of certain psychological mindsets: immersion (Brown and Cairns, 2004, Ermi and Mäyrä, 2005), presence (Takatalo et al., 2006), and flow (Kivikangas, 2006, Polaine, 2005, Sweetser and Wyeth, 2005). In the study presented here, we are looking at combat experiences in levels of the First-Person Shooter (FPS) game Half-Life 2 (Valve Corporation, 2004).

2. LEVEL DESIGN CONSIDERATIONS

We designed a set of stimuli levels with the help of the Half-Life 2 Source SDK (Valve Corporation). The design was focused on creating different player experiences, especially boredom, immersion and flow (Stellmach, 2007). All levels were played in the following order:

1. Secret Corridors (Immersion)
2. Church Walk (Boredom)
3. Flow Check (Flow)

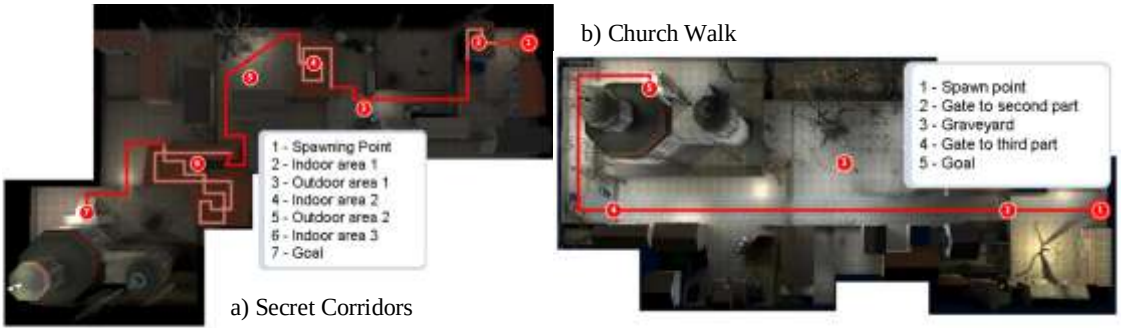
The names of the levels were merely codenames for internal production. When presented to the participants of the experiment, we left out the descriptive names and just called the levels experiment 1-3.

However, for this report on the design of the levels, we find the descriptive names better to refer to the different level designs.

2.1.1 Secret Corridors

Players must reach a target destination, the church, in this level. The way to the target leads through three indoor areas and four outdoor areas (see Figure 1, left side) designed in the style of the Half-Life 2 level "Ravenholm". Although the level layout is linear, it was designed so combat areas would be followed by resting areas, where the player could resupply and explore the scenery. At the start enemy strength is weaker than close to the end. Alternating between combat and exploration, the level focuses on creating player immersion. Estimated playing time for this level was 10 minutes (depending on player skill).

Figure 1.Bird's eye view of Secret Corridors (left) and Church Walk (right). Secret Corridors was designed for feeling immersed, Church Walk for feeling bored



a) Secret Corridors

b) Church Walk

2.1.2 Church Walk

This level was designed to evoke a feeling of boredom in reasonably skilled FPS players. Set design was sparse and level layout linear. The player needed to walk to a church, almost following a straight line (see Figure 1, right side). Estimated playing time for this level was 6 minutes. The level was designed as an outdoor-only environment, using the "Ravenholm" setting. Enemies are of the same type (slow melee attackers), scripted to cause less attack damage to keep challenge low. We expected combat encounters to bore players due the repetitive nature and low frequency.

2.1.3 Flow Check

Standing apart from the other levels, this one uses an indoor-only futuristic setting. The player must walk through three rooms connected with doors that only open after the player has killed all enemies in each room. The first room includes low challenge combat against slow melee enemies, while in the second room combat frequency is higher and enemy types are more aggressive. In the last room, combat is very challenging against small, fast and aggressive enemies, actually making survival possible only by evading the room using a ladder. We added challenge to the attack process as the player has only a melee (crowbar) and a sniper weapon (crossbow) with limited ammunition and slow reload times. We wanted to explore gradual challenge and whether that engages players in a flow state.

3. METHODS

3.1 Subjects

Data was recorded from 25 healthy male higher education students, aged between 19 and 38 (*Mean*= 23.48, *SD* = 4.76) years. As part of the experimental setup, demographic data was collected with special regard to the suggestions made by Appelman (2007). Most of the participants were right-handed (88%). All the participants owned a personal computer (PC) and 96% rated this as their preferred gaming platform (out of multiple choices), followed by Xbox 360 (56%), Playstation 3 (52%) and PS2 (48%). More than half of all participants said they would buy games more than once a year (68%) and all of them played games at least

twice a week. Most (84%) played between two and four hours a day. The preferred mode of play was console single player (44%) or pc multiplayer (36%). More than a third (36%) rated First-Person Shooters (FPS) as their favorite game type. All the participants started playing digital games before they were twelve years old. None of the subjects received any compensation for their participation.

3.2 Measurement instrumentation

For this study, we used a variety of self-report measurements to acquire demographic and psychographic data of players. We used many psychological questionnaires (Buss and Perry, 1992, Carver and White, 1994, Eysenck et al., 1985) to measure behavioral displacements and evaluative gameplay self-reports with the game experience questionnaire [GEQ] (IJsselsteijn et al., 2008). While we used the results reported in the original publications as a baseline, the development of the GEQ is still ongoing so that we expect to be able to compare our measures (regarding such a baseline) once the first study has been published.

3.3 Procedure

We conducted all experiments on weekdays in the Game and Media Arts Laboratory with the first time slot beginning at 10:00h and the last ending at 20:00h. General time for one experimental session was 2 hours. Each participant was invited to the lab and briefed about the experimental procedure. The participants needed to fill in an informed consent form and a photographic release form. Next, participants filled in the questionnaires. Then, the participants played the game levels. Each game session was set to 10 minutes, but in general participants finished quicker. After each level, each participant was given a paper version of the game experience questionnaire to rate their experience (IJsselsteijn et al., 2008). After completion of the last questionnaire, participants were thanked for their participation and escorted out of the lab.

4. RESULTS

The results of the short scale EPQ-R were scored according to Eysenck et al. (1985). Interestingly, the most reliable of the components, extraversion ($\alpha = 0.86$) scores also the highest mean (7.8). Thus, we can say that most of the participants in our experiments have extrovert personality. The mean values for psychoticism, neuroticism and lie are all close to expected low values for that age range. The results from BIS/BAS scales were as expected with two exceptions. Two BAS mean scores are rather low, indicating stronger reward responsiveness (1.6) and a more fun seeking (2.0) persona. In general, total mean scores are lower than those reported by Carver and White (1994). The results from the aggression questionnaire were similar to those reported by Buss and Perry (1992) with the exception of the hostility mean score, which was lower (18.9) than that reported for men in the original report.

Although, game experience was assessed independently for each level, the order in which the levels were played might have impacted the results. Our personal observations showed that players with less experience in playing Half-Life 2 liked this level better than those that played the game before. Church Walk scored highest in the areas of competence and positive affect (see Table 1). They also proved to be the most reliable of the measures ($\alpha = 0.92$). Church walk scored low for the flow component (1.6). Even though negative affect scores low (1.2) as well, it has the widest range (1.5). This could be interpreted as different influences of the challenge level on different kinds of players. Players that played Half-Life 2 before might not have liked the overall impression of Church Walk, which could also be related to a low score on sensory and imaginative immersion (0.9). This could also be the impact of having played Secret Corridors before. The internal consistencies of the items on the game experience questionnaire were found to be good with α ranging from 0.75 to 0.92 for this level.

This was not the case for the next level, Secret Corridors, where the reliability of items on the GEQ was unfortunately found to be barely acceptable for challenge ($\alpha = 0.53$) and tension ($\alpha = 0.67$) components (see Table 1). The other items scored well in the range between 0.79 and 0.86. Again, competence was found to have the highest range between values (range = 1.7). In general, we could find higher ranges between the values for this level. The lowest range was luckily for the measure of game immersion, which is exactly what this level was designed for (0.9).

Table 1. GEQ statistics for all three levels, N = 24, N of Items = 6, GEQ scale ranges from 0 to 4

GEQ components	Cronbach's α Church Walk	Cronbach's α Secret Corridors	Cronbach's α Flow Check	Means Church Walk	Means Secret Corridors	Means Flow Check
Immersion	0.87	0.79	0.86	0.85	1.37	1.21
Tension	0.82	0.67	0.88	1.03	0.83	1.90
Competence	0.81	0.86	0.94	2.55	2.50	1.99
Flow	0.92	0.84	0.94	1.61	1.94	2.21
Negative affect	0.85	0.80	0.89	1.18	0.71	0.89
Positive affect	0.92	0.86	0.91	2.08	2.60	2.28
Challenge	0.75	0.53	0.68	0.82	1.27	2.34

This level was designed for a more immersive experience, which shows in the highest mean (1.4) for all three levels for the item sensory and imaginative immersion. We must assume that none of the levels managed to score excellent on immersion since all scores are below the median. However, this could also be because of the experimental conditions of playing in a laboratory rather than in an ecologically valid gaming environment.

The highest means are found for the items competence (2.5) and positive affect (2.6), which are also found to be the most reliable of the measures ($\alpha = 0.86$). The competence level was surprisingly similar to that of the easy level even though this level has more intense combat challenges and a higher enemy count. We could argue that this reflects in the higher values of flow (1.9) and challenge (1.3). It could also be argued the design of the level with ebb and flow of combat challenges (scattering relief areas inside the level) led to this and to higher immersion. Thus, also positive affect could be connected to immersion and challenge.

In the last level of our study, Flow Check, the reliability of the challenge item was found to be only just acceptable ($\alpha = 0.68$), but all other measures worked reliable, α ranging from 0.86 to 0.94 (see Table 1). As expected this level rated highest on the challenge item (2.3). This also worked well for flow as this higher level of challenge led to players experiencing a flow state (2.2). Surprising was that even with an increase in challenge, players still felt competent enough, with competence being just at the median of the GEQ (2.0). Positive affect was found to be still high (2.3), but lower than for Secret Corridors. Thus, positive affect could likely be connected to immersion. Tension scored highest in this level (1.9), suggesting that tension and challenge are interrelated.

Last, we found that standard deviation on the easy level was low for all items (0.2 to 0.6), a little higher on the medium level (1.0 to 1.2) and highest on the challenging level (1.3 to 1.4). This could explain that individual differences are more visible the higher the challenge of the gameplay is. However, to truly find out about this, a study that looks at the interrelations of the different levels would need to be conducted in the future.

5. CONCLUSION

The preceding section presented the results of the Game Experience Questionnaire (GEQ), which we used to provide added construct validity of the GEQ. From a design aspect, we created the three levels with different aspect of the game experience in mind, low challenge and boredom (Church Walk), immersion (Secret Corridors) and high challenge in combination with combat-induced flow (Flow Check).

While in the first two levels, the player could choose among various weapons, the last level was explicitly designed only around one shooting weapon (crossbow). The amount of ammunition was counted to be just enough if the player shot each enemy with one arrow. Of course, this did not happen often because of the frequency of attacks and players trying to find a safe spot in the room. If ammunition was out, the players had to fall back to the melee weapon, the crowbar, and jump around while trying to fight off the enemies. The attack frequency was raised throughout the level, which would naturally also raise the challenge level. This was the only level where players could die more than once and explicitly because they were killed by the enemies. What is interesting about this is that one would expect a higher level of frustration and maybe negative affect than that revealed by the experiment.

The results above show that the last level is not the one with the highest negative affect mean. The low challenge and boredom level is the one that scores highest on negative affect, thus we could say the absence of enough challenge leaves the players in a negative emotional state, while more challenge provides a much better experience in general, rating high on positive affect and flow. This may also be associated with an unusual aspect of the level design, compared to regular Half-Life 2 gameplay, concentrating on fighting and managing resources (ammunition); this reduces general Half-Life 2 gameplay to the minimal aspect of mastering different forms of combat.

ACKNOWLEDGEMENTS

This experimental pilot study was supported by the European Community FP6 NEST project "The Fun of Gaming: Measuring the Human Experience of Media Enjoyment". We would like to thank project partners and Wouter van den Hoogen, Anders Tychsen, Niklas Ravaja, Matias Kivikangas, Dennis Sasse, and the staff of the BTH Game and Media Arts Laboratory, especially Sophie Stellmach, for their supportive feedback during conception and analysis.

REFERENCES

APPELMAN, R. (2007) Experiential modes of game play. *Situated Play, Proceedings of the DiGRA 2007 Conference.* Tokyo, Japan, DiGRA.

BROWN, E. & CAIRNS, P. (2004) A Grounded Investigation of Game Immersion. *Conference on Human Factors in Computing Systems*. Vienna, Austria, ACM Press.

BUSS, A. H. & PERRY, M. P. (1992) The aggression questionnaire. *Journal of Personality and Social Psychology,* 63**,** 452-459.

CARVER, C. S. & WHITE, T. L. (1994) Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology,* 67**,** 319-333.

ERMI, L. & MÄYRÄ, F. (2005) Fundamental Components of the Gameplay Experience: Analysing Immersion. *DiGRA conference Changing views: worlds in play.* Vancouver, Canada, DiGRA.

EYSENCK, S. B. G., EYSENCK, H. J. & BARRETT, P. (1985) A revised version of the psychoticism scale. *Personality and Individual Differences*, 6**,** 21-29.

GRIFFIN, S. N. (2005) Push. Play: An Examination of the Gameplay Button. *DiGRA 2005 Conference: Changing Views – Worlds in Play.* Vancouver, Canada, University of Vancouver.

IJSSELSTEIJN, W., DE KORT, Y., POELS, K., JURGELIONIS, A. & BELLOTTI, F. (2007) Characterising and Measuring User Experiences in Digital Games. *International Conference on Advances in Computer Entertainment Technology.* Salzburg, Austria, ACM Press.

IJSSELSTEIJN, W. A., POELS, K. & DE KORT, Y. A. W. (2008) Measuring player experiences in digital games. Development of the Game Experience Questionnaire (GEQ). *Manuscript in preparation.*

KIVIKANGAS, J. M. (2006) Psychophysiology of flow experience: An explorative study. *Faculty of Behavioural Sciences, Department of Psychology.* Helsinki, Finland, University of Helsinki.

POLAINE, A. (2005) The flow principle in interactivity. *Proceedings of the second Australasian conference on Interactive entertainment.* Sydney, Australia, Creativity & Cognition Studios Press.

STELLMACH, S. (2007) A psychophysiological logging system for a digital game modification. *Department of Simulation and Graphics.* Magdeburg, Otto-von-Guericke-University.

SWEETSER, P. & WYETH, P. (2005) GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)*, 3**,** 3-3.

TAKATALO, J., HÄKKINEN, J., KOMULAINEN, J., SÄRKELÄ, H. & NYMAN, G. (2006) Involvement and presence in digital gaming. *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles.* Oslo, Norway, ACM Press.