

Using Heart Rate and Machine Learning for VR Horror Game Personalization

1st Sumaira Erum Zaib
Department of Computer Science
Tokyo Institute of Technology
Yokohama, Japan
erum.s.aa@m.titech.ac.jp

2nd Masayuki Yamamura
Department of Computer Science
Tokyo Institute of Technology
Yokohama, Japan
my@c.titech.ac.jp

Abstract—In this paper, we explore if we can personalize horror games in Virtual Reality using auditory and visual stimuli with the help of different machine learning algorithms. Based on the heart rate of the subjects we personalize the sound and lighting effects of the game in two different environmental settings. Gradient Boosted Tree Regression, Random Forest Regression, and Tree Ensemble Regression were used to predict which sound and lighting effects should be used in subsequent levels to increase the horror aspect. In order to have a realistic game experience and due to the ongoing coronavirus pandemic, participants were recruited online. Participants could play the game wherever and whenever they wanted. The participants were also asked to complete Self-Assessment Manikin tests after playing the game. We present our discussions and observations of how different factors affect the heart rate in the game and if the heart rate data aligns with the participant’s Self-Assessment Manikin test data.

Index Terms—Horror, Machine Learning, Virtual Reality, Affective Gaming, Heart Rate

I. INTRODUCTION

Video games are a great way of story telling and creating an experience that a player feels invested in. People play games to vent out their frustrations, relieve stress or to explore different emotions [8]. Video games offer a way to experience different emotions and deal with them safely [1], [2]. In the past two decades, there has been a surge in horror games and attempting to make them as immersive as possible.

Games like *Silent Hill: Shattered Memories* (2009) [35] and *Until Dawn* (2015) [36] perform psychological tests on the player. The story-line then changes according to the player’s answers. *Amnesia: The Dark Descent* (2010) [24] inflicted psychological horror using distorted visuals, different sounds and made the player question logic slowly. The limitation of these type of games is that there is no user-specific personalization. With replays players can win these games but lose the horror aspect [37].

Adapting the game to the player’s emotions and physiological states is known as affective gaming [31]. With the increase in immersive games, horror game designers and developers need to carefully control the user experience to induce proper amount of emotions [25], [33] such as fear and anxiety. This should be done while making sure the player is not overwhelmed or underwhelmed. Efforts have been made to analyse and predict emotions during game play [15]. However,

to achieve that, the player has to go through questionnaires and psychological tests.

Current researches that induce emotions, based on Electro-Dermal Activity (EDA), Respiratory Rate (RR), or Heart Rate (HR) [3], [11], [30], [49], use special sensors for biofeedback to personalize games and experiences. Tests like Self-Assessment Manikin (SAM) [5] are usually administered in these studies to evaluate if the observed results match with the experience of the subject. There have not been many researches pertaining to the real world scenario where game is played wherever and whenever the player wants. Individual emotional differences of the players are also usually not taken into consideration in such studies.

The purpose of this study is to find out if using just the heart rate from player’s own smart watch to modify the horror game is enough to induce fear and to increase the horror aspect of the game. We also try to understand if playing different types of environment (indoor and outdoor) has any effect on the players. This will help horror game developers to incorporate heart rate into games for personalization and for better control of game tension based on environmental changes. For this study, we developed a Virtual Reality (VR) horror game for greater immersive experience. The game was personalized using machine learning algorithms and heart rate data. Machine Learning (ML) algorithms namely Gradient Boosted Tree Regression, Random Forest Regression, and Tree Ensemble Regression were used to create personalized prediction models for each participant. The game was played three times by the participants, with each session lasting for ten minutes maximum.

This research, using HR, aims to answer the following

- Does Valence decrease for ML-boosted scene?
- Does Arousal Increase for ML-boosted scene?
- Does HR correlate with SAM results?
- Does HR differ after replays?
- How does environment and adaptability effect HR?

The rest of the paper is structured as follows. Section II provides important background information for understanding horror games and existing research related to horror games. Experimental setup is elaborated in Section III and our game design is explained in Section IV. Section V presents the results of the experiment and in Section VI the findings of

the experiment are elaborated. The paper ends with conclusion and future direction in Section VII.

II. RELATED WORK

Although horror genre creates negative emotions, its popularity cannot be denied [4]. People are fascinated with fear because it is different from daily life [9] and can be experienced safely using horror movies and games.

A. Emotions

How people feel different emotions is a very personal and individual experience [55]. It differs significantly from person to person based on their gender, culture, life experiences and personality [16]. Research has shown that men and women have different coping mechanism when dealing with stress and other negative emotions [27]. Similarly, how people react and consume horror media depends on their personality, age, gender and beliefs [14].

Using customised Lightstone device, Dekker and Champion [11] used Heart Rate Variability (HRV) and skin response data to enhance the horror game's volume and the player's moving speed. If the player could control their emotional state they could manipulate the game further. The results showed that players preferred biofeedback enhanced game over non-biofeedback enhanced game. However, no quantitative results were specified in this study.

B. Audio Visual Stimuli

In order to create a personal experience in horror games audio visual stimuli should be adapted according to the player. Some people may be more affected by visual stimuli and some more with audio stimuli. Some people might need a combination of both to experience fear [22]. Graja and Lopez [22] built a game using finite state machines that make use of audio and visual effects to study the impact of Horror Games on Galvanic skin Response. Sounds were reported to be the most important factor to induce greater stress and anxiety in this research. Different frequencies, amplitudes of sounds can also produce uncanny feeling [20], [21], [23], [38]. [50] shows that audio is very important to increase immersion in games.

C. Environment

Environment plays an important role in creating an immersive horror game experience. Different lighting and sounds contribute in making the player feel vulnerable [53]. Open spaces or closed spaces trigger different kinds of emotions for people, more so if the player has claustrophobia [32], [43].

Maintaining an element of obscurity by limiting the vision of the player also helps in making the player feel vulnerable and unsafe from potential dangers. Obscurity can be created not only visually but also through sounds. Sudden noises or complete silence will trigger negative emotions in the player and make them uncomfortable [22], [43]. It can be imagined that having a familiar environment would likely not trigger fear and vulnerability. However, if unusual and unexpected things happen in a familiar setting it can make the player uncomfortable and question their safety in the game [28].

D. Immersion

Immersion is a feeling described as being present in the game while losing awareness of real time and place [26]. Currently out of all game formats, Virtual Reality (VR) provides the most immersive experience to the player [45]. It creates an illusion of actually being present in the game. Real time movement in the game make the experience even more close to reality [44]. Due to this, VR is used extensively in emotion research. Inducing different kinds of emotions creates real life reactions and helps the researchers in understanding different emotions better [53].

Madsen [30] did a comparative study of horror game players and watchers using electrodermal activity (EDA), respiratory rate (RR), heart rate (HR) and self reported fear data. Results showed players had greater variance in physiological data than watchers but no significant difference was noticed in self reported fear data.

Slater [44] explained that in VR Place illusion (PI) is experiencing an illusion of being in a place even though the player surely knows that it is not true. PI corresponds to environmental and audio visual cues. Slater [44] also explained that Plausibility illusion (PSI) is experiencing something is real even though the player knows it is surely not real. PSI corresponds to cues that prompt action such as enemy attacking. Based on this, Jih-Hsuan and Lin [29] categorized different fear elements in a zombie survival game and studied different coping strategies employed by players. Their results showed that PSI elements cause more stress and fear in players. [19] did a preliminary study to detect fear when playing games using electroencephalography. [34] used player metrics to personalize horror game with promising results.

E. Heart Rate and Heart Rate Variability

Heart Rate Variability (HRV) [42] is the time fluctuations between consecutive heart beats. HRV can be used to give acceptable indication of emotional changes when the stimuli is strong [13]. Low HRV corresponds to less variation in time intervals between heart beats, meaning regular heart beat and a state of calmness. High HRV shows more variance and thus irregular heartbeats, correlating to the state of excitement and short term stress [41], [56]. [52] showed that HR is a good indicator for horror game's affectiveness.

Araki et al. [3] claimed to increase the pulse rate of the players by routing players according to the increase in their HR. However, they did not provide results for this claim.

Sun, Lin and Lu [47] used eye movement to analyse fixation and saccade of the players during horror game play. HR and pupil changes were shown to be a good measure to get immersion levels of the player. A decline in fixation frequency and length was also noticed when players played longer and got more immersed. However, these claims were not backed by results in the paper.

Ueoka and Ishigaki [49] developed a horror emotion amplification system using biofeedback. The experiment used two types of pseudo HR, one that adapts to the subject's heart in

real time and the other that increases in step-wise manner. However, the results were not conclusive.

Voigt-Antons et al [53] studied if different emotional states can be induced in players after playing different types of content with different levels of interaction in VR games. SAM scores were used to analyse valence and arousal in the players. HR variability was used as the main measure to support the findings physiologically. The results indicate that playing horror game showed an increase in arousal and a decrease in valence which indicated fear in SAM. Following the horror game, when nature experience was played, the SAM scores indicated decrease in arousal and increase in valence.

At present, all research related to emotions and horror that involve HR uses different HRV parameters for evaluation and analysis purposes. To measure HRV, special gadgets are tied to the chest or wrist of the subject. However, this is not something people usually have and designing games around these gadgets will result in a very limited inaccessible game. Therefore, in this study we investigate whether smart watches that measure HR can be used to collect data and personalize horror game.

F. Machine Learning Algorithms

Ensemble learning is used to achieve high predictive performance using multiple learner models and combining their predictions. Ensemble learning also reduces the risk of overfitting in small data sets [39], which is perfect for our small data for every subject. The data for the player in our experiment will be updated after each game play, requiring the model to re-train after every game session. The resulting model could then have higher or lower prediction accuracy than the first model. Hence, to improve prediction accuracy, three different ensemble models were trained at run-time. The model with the highest accuracy was then selected for personalization settings.

We use three different ensemble machine learning algorithms namely Gradient Boosted Tree Regression, Random Forest Regression, and Tree Ensemble Regression. The specifics of these algorithms are discussed in subsection IV-G. This section provides a basic overview of these algorithms.

Gradient boosting [17] works by building multiple models and optimizing a differential loss function. Gradient boosting combines multiple weak models into a single more accurate and robust model. Random Forest Regression [6] is a supervised machine learning algorithm that creates a single model by using predictions from ensembles of different machine learning algorithms. Lastly, Tree Ensemble Regression [40] uses multiple weighted regression trees to make a more accurate prediction. These algorithms can be used for both classification and regression.

Vachiratamporn et al. [51] used electroencephalography (EEG) and electrocardiography (EKG) signals, and keyboard–mouse activities data to create prediction models for the affective state of the players before and after experiencing a horror event in a survival horror game. Their results showed increase in game enjoyments after multiple game rounds but also a decrease in fear.

III. EXPERIMENT SETUP

A. Participants

This study was approved by the ethics committee of the relevant faculty. The experiment conducted for this research was not done in laboratory setting. There were two reasons for that. First, due to the ongoing Coronavirus pandemic recruiting people online and having them play the game in their own environment was safer for both the participants and the researchers. This also allowed us to have a diverse group of participants. Secondly, participants playing at home are more engaged and immersed in the game than in laboratory [48].

Generally people do not play games in special locations with special sensors to detect different physiological responses to the game. People play games whenever and however they want. They might play standing, sitting, with headphones, without headphones and so on. Therefore, it is difficult for game developers to practically translate horror game research into their games. Considering VR games require proper set up and play area, it would always be preferable to play at home. Therefore, for this research, participants were recruited online on Reddit and Facebook VR communities. They were required to already own all the devices needed for this research.

20 people (10 Males, 10 Females) participated in this research with ages from 22 years old to 38 years old. The participants were used to playing VR games. The game was shared with the participants after completing the consent form and confirming the ownership of the required devices. No incentive was given for participating in this research. Participants were free to quit the research at any moment.

B. Self-Assessment Manikin (SAM)

Self-Assessment Manikin (SAM) [5] is a pictorial questionnaire to determine emotion using valence, arousal and domination. Low valence and high arousal corresponds to horror and fear [53]. In our study, a nine point SAM questionnaire was filled by the participants after every session to collect valence and arousal values.

After each game play session, the participants were required to fill a 9 point SAM questionnaire about their pleasantness feeling (valence) and state of calmness (arousal) during the game. 9 represented most pleasant and least calm. 1 represented most unpleasant and most calm.

C. Devices

Participants were required to own an Oculus Quest or Oculus Quest 2, an android Smartphone (Minimum Android version: Oreo), a smart watch to detect HR and sync to Google Fit. The participants were instructed to turn on Continuous Heart Rate monitoring setting on their smart watch that allows HR capture without any activity specification.

D. Experiment Flow

After installation of VR game on headset and mobile application on android phone, the following steps needed to be repeated at least three times while wearing smart watch

- Play 10 minute Scene (Forest or Hospital)

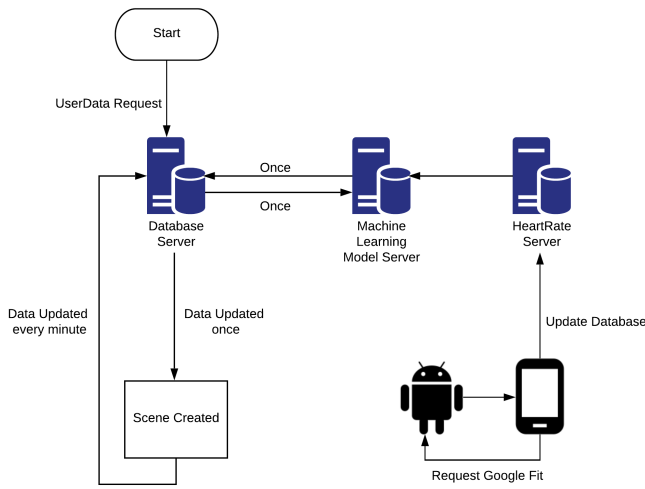


Fig. 1. Communication of data between the VR game, servers and android mobile application.

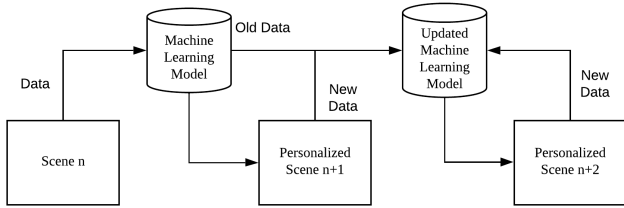


Fig. 2. Simple flow of how the game is personalized.

- Fill out SAM questionnaire in the mobile application
- Upload HR data using Mobile Application

Figure 1 shows the communication of these systems.

IV. GAME DESIGN

To collect the required data for the experiment, we created a VR Horror Game using Unity Engine. Participants were required to play the game at least three times for 10 minutes maximum. The session could end early if the goal of the scene was reached. The light and sounds setting in the first session were generated randomly. For subsequent sessions, the setting values were retrieved from machine learning models (Figure 2). To create a realistic experience, no jump scares or enemies were used in the game.

A. Scenes

To explore how different environments effect individuals two types of scenes were created. The scenes were modified so that the player could explore them with sounds and lights adapted according to the player's HR.

- An outdoor forest scene (Figure 3), with abandoned houses and buildings that the player could go inside, was designed using Flooded Grounds unity asset. On start, a

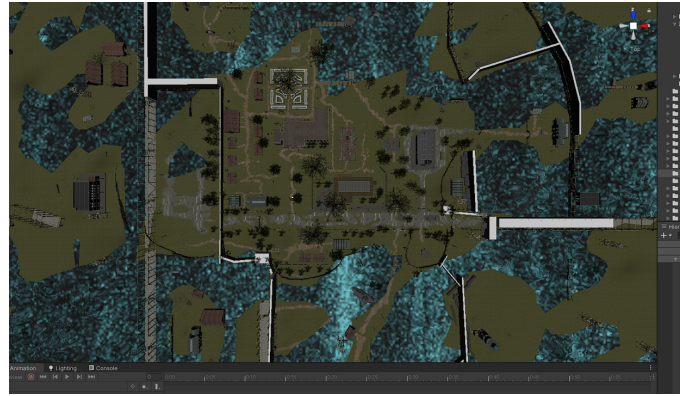


Fig. 3. Map of the forest used in this experiment.

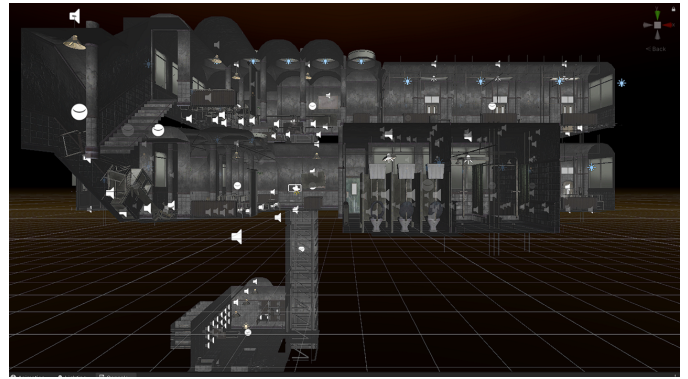


Fig. 4. Model of a three floor hospital used in this experiment.

random door was assigned to be the glowing exit door. The goal of the forest scene was to find a glowing door.

- An indoor three floor hospital scene (Figure 4) was designed using HE - Abandoned Hospital v.1 unity asset. On start, a key was placed randomly in the hospital. The goal of this scene was to find the key and exit the main door. To encourage exploration and player's attentiveness of surroundings, the key could be on the ground, inside drawers or on top of any object.

B. Sounds

At any moment in the game there were four sounds playing. One background sound and three sounds that surrounded and moved with the player. The three sounds were placed at different distances from the player and overlapped each other.

Background sound was selected from four categories: Wind, Rain, Insects and Others. The other three sounds were selected from 19 different categories namely: Crow, Dog, Frog, Others, Man_Crying, Woman_Crying, Baby_Crying, Bell, Ground, Footsteps_indoor, Footsteps_outdoor, Knocking_iron (knocking sounds on iron door), Knocking_wooden (knocking sounds on wooden door), Breathing_monster, Eating_monster, Heartbeat_monster, Laughing_monster, Ear_monster and Mouth_monster. In total, 187 sounds from 23 different categories were used in this game.

C. Lights

Since both the scenes had indoor spaces, different types of lamps were used in both the scenes. In both scenes the intensity of lamp light could be on, off or flickering. Also, the lamps could turn on or off automatically around the player.

D. Controls

Oculus Quest was the required VR Headset for this experiment. Controller joysticks were used for navigation. The player was equipped with a torch in left hand. The intensity of torch light could be controlled using the left grip button. The right grip button was used to interact with doors, drawers, cupboards and hold the key in the game.

E. Mobile Application

An android mobile application was designed to collect HR data in beats per minute (bpm) from Google Fit. Therefore, players with any smart watch that could sync with google fit could play the game. The participants were required to play the game at least three times and sync data with Google Fit after every session. Then the provided mobile application would upload their HR data to our servers. This was also designed after recommendation from the ethics committee, so that the participants could take a little break after playing one session and prevent possible VR sickness [12].

F. Game Data Collection

10 participants were required to play the outdoor (forest) scene first and the rest had to play indoor (hospital) scene first. Each scene was played at least once and players could choose the third scene on their own. A remote server was created to collect data from the game. After every minute in the game the sounds and lights would change and the updated values would be added to the database. The heart data collected from Google Fit provided HR for every minute. Therefore, a 10 minute session resulted in at least 10 HR data points. HR data was then matched with the game data using timestamp date, hour and minute values. The data prepared was then used as training data for prediction models in subsection IV-G. Session 2 would be personalized based on data collected from session 1. Session 3 would be personalized based on data from session 1 and session 2 (Figure 2).

G. Personalization through Machine Learning

To personalize the game for increased horror aspect, we wanted to increase the HR of the player when the sound and light setting changed every minute. Throughout the minute, the player would get used to the sounds and lights.

For this purpose, regression was performed using three different algorithms with HR as target variable. As explained in subsection II-F, three ensemble models were used. Since the model is being re-trained after every session, the accuracy of the new model could be higher or lower than the previous model. Therefore, to improve chances of good accuracy, three different regression algorithms were used. Considering that the model is personalized and the training data size is not huge

(less than 1000 rows for each participant), it was feasible to do model training at run time at the start of game session.

Based on the data collected in Section IV-F variables prepared for the models were: Volume (Value from 0 to 1) Pitch (Value from -3 to 3), StereoPan (Value from -1 to 1), Sound (Value from 1 to 187), Sound_Category (Value from 1 to 23), Light_Value (Value from 0,1,2), Heart_Rate (Value from 50 and above). Light value 0 corresponds to off, 1 corresponds to on and 2 corresponds to flickering light.

At the start of every session Random Forest (Models: 100, Static Random Seed), Gradient Boosted Tree (Models: 100, Learning Rate: 0.1, Alpha: 0.95, XGBoost missing value, Static Random Seed), Tree Ensemble (Models: 100, Fraction of Data for single model: 1, Static Random Seed) regression models were trained with 80 percent data and tested with 20 percent data using linear sampling.

To initialize new session of the game with different sound and light setting, 500 rows of random data was generated without the HR variable. The model with the highest accuracy was chosen at run-time and was used to predict the HR of the randomly generated test data. The data was then sorted from highest to lowest HR. The 40 rows with highest HR were fed to the scene to initialize the game and change the settings of 4 sounds and lights every minute.

V. RESULT

13 out of 20 participants played all three sessions for 10 minutes whereas others ended sessions early by completing the goal. On average, session 1 was played for 9.85 minutes, session 2 for 9.75 minutes and session 3 for 9.35 minutes. Every session was played for at least 5 minutes.

A Wilcoxon-Signed-Rank Test [54] was done to check if the same scene when ML-boosted had low valence, high arousal and variation in HR. Valence (Statistic = 2, $p = 0.009$) and arousal (Statistic = 14, $p = 0.048$) showed statistically significant result but average HR (statistic = 31, $p = 0.5$) and HR variance (statistic = 40, $p = 0.7$) did not. To check environment adaptability 10 participants played the forest setting first and 10 participants played the hospital setting first. The mean and variance of SAM results and game HR data can be seen in Figure 5, Figure 6 and Figure 7 respectively.

Friedman Test [18] was done to check if over the three sessions there was a decrease in valence, increase in arousal and variation in HR. The results for valence and arousal for different groups are shown in Table II. Average HR and HR variance did not produce significant results for any group. Therefore, they are not presented here.

For individual subject analysis, Mann-Kendall test [46] was done to check for trends in individual subject HR over three sessions. A high positive S value shows an increasing trend. A low negative S value shows a decreasing trend. 9 subjects had statistically significant trend. An increase in HR was observed for 13 subjects and decrease in HR for 7 subjects (Table I).

To understand gender differences, correlation was checked Correlation of Male subjects' HR

- Variance with SAM Valence: 0.47 and Arousal: -0.42

TABLE I
MANN-KENDALL TEST FOR TREND DETECTION (FEMALE SUBJECTS: 1 TO 10, MALE SUBJECTS: 11 TO 20). * INDICATES SIGNIFICANT RESULTS.

Subject	Minimum HR	Maximum HR	Mean HR	Standard Deviation	Kendall's tau	S	Var(S)	p
1	58	97	72.419	13.436	0.580	265	3442.333	<0.001*
2	59	92	78.679	7.912	-0.318	-117	2531.667	0.020*
3	51	82	71.000	7.710	0.284	90	2040.000	0.046*
4	62	93	79.182	7.904	-0.183	-94	4131.333	0.144
5	50	93	75.313	9.509	-0.407	-198	3778.000	0.001*
6	53	91	79.484	9.003	0.088	40	3440.667	0.495
7	71	98	86.677	6.300	-0.140	-64	3440.667	0.275
8	64	90	79.182	4.799	-0.088	-44	4062.667	0.490
9	62	94	82.462	7.527	-0.094	-30	2048.000	0.507
10	50	94	83.778	10.551	0.050	17	2279.667	0.722
11	77	99	89.133	5.981	-0.265	-112	3106	0.044*
12	78	109	88.000	7.102	0.042	22	4142	0.744
13	52	98	76.000	9.619	0.313	186	4938	0.008*
14	78	111	91.000	7.056	0.117	62	4142	0.343
15	62	114	87.000	15.856	0.656	305	3455.667	< 0.001*
16	65	94	79.281	7.292	0.270	132	3787.333	0.032*
17	72	98	84.909	7.547	0.033	17	4127	0.791
18	51	92	74.969	8.185	-0.130	-63	3775	0.305
19	51	96	79.875	10.954	0.155	76	3786.667	0.217
20	54	95	82.333	7.825	0.297	153	4131	0.017*

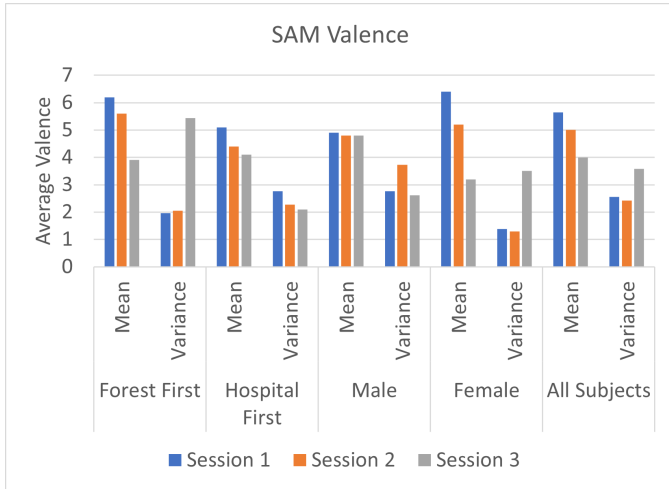


Fig. 5. SAM Valence: Forest first, hospital first, Male, Female and all subjects.

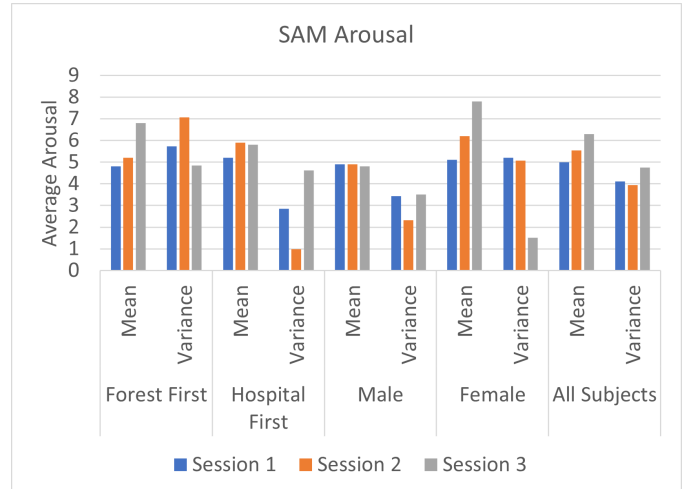


Fig. 6. SAM Arousal: Forest first, hospital first, Male, Female and all subjects.

- Average with SAM Valence: 0.29 and Arousal: -0.68 Correlation of Female subjects' HR
- Variance with SAM Valence: 0.92 and Arousal: -0.95
- Average with SAM Valence: -0.11 and Arousal: 0.15

VI. DISCUSSION

Going back to the questions raised in section I, we can analyze the results in section V to answer these questions.

Does Valence decrease for ML-boosted scene? Yes. This shows that the game was able to increase the unpleasant feeling using machine learning over different sessions.

Does Arousal Increase for ML-boosted scene? Yes, throughout the sessions we see a steady increase in average arousal.

Does HR correlate with SAM results? For female participants HR variance correlated strongly with SAM Valence and arousal. However, strong correlation was not observed for

TABLE II
FRIEDMAN TEST RESULTS (* INDICATE SIGNIFICANT RESULTS.)

	Valence		Arousal	
	Statistic	p	Statistic	p
Forest First	8.36	0.01*	4.2	0.12*
Hospital First	0.08	0.048*	5.06	0.08*
Male	0.87	0.65	0.53	0.77
Female	15.23	0.0005*	11.05	0.004*
All	12.09	0.002*	8.34	0.01*

male participants. [16], [27] reports that although men and women experience emotions equally, women tend to be more expressive about them. This could be why self-reported results for female participants correlate with HR variance much more than male participants. Table I also shows that 8 out of 10 males

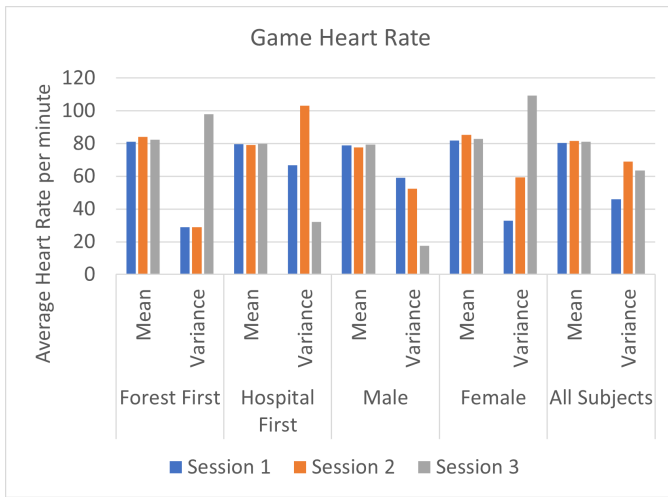


Fig. 7. HR: Forest first, hospital first, Male, Female and all subjects.

showed increasing trend in their HR whereas for females it was an equal split for increasing and decreasing trend.

Does HR differ after replays? Average HR does not differ much from session 1 to session 3. For interactive task based games replay value increases due to better flow of the game with exposure and enjoyment remains stable [37]. HR variance decreased for male subjects and increased for female participants. However, this change was not statistically significant.

How does environment adaptability effect HR? People who played hospital first had an increase in HR variance when they played forest. Whereas people who played forest first did not have a huge increase in variance when they played hospital. Studies [53], [7], [10] show that exposure to outdoor nature setting helps people to calm down. In the third session, 13 out of 20 subjects chose to play the forest scene. Preference for open outdoor environment was observed in participants. We see in Figure 7 that for people who played indoor (hospital) first and then continuously outdoor (forest) had much lower variance in their HR. Although average HR did not change, playing a nature setting after an indoor setting still resulted in low average valence and high average arousal along with high HR variance. This shows that even after the player had experienced both environments, horror aspect still increased.

Based on the results, we can see that horror experience is very different for different individuals and for at least female participants HR variance calculated from player's own smart watch can be used by game developers. However, more research is required to get more reliable results for male participants. Combinations of indoor and outdoor experience can also be used to control the horror experience.

VII. CONCLUSION AND FUTURE WORKS

In this research, we observe that current horror games require interference with questionnaires and psychoanalysis of the player to change the routes of the game. Current efforts to incorporate biofeedback into games have only been tested using special sensors in laboratory settings.

However, in our study we allow the user to play a horror game in their own space. We developed a custom VR horror game that would personalize audio and light settings after each session for the participant based on their HR. In our results, we show that HR captured from the player's own smart watch showed significant increase in HR for 6 subjects and significant decrease in HR for 3 subjects. The results indicate that horror experience is personal and not the same for everyone, even when exploring similar environments.

Based on SAM results and HR variance correlation game developers can incorporate HR into game design for personalization of horror game experience. We also see how environmental adaptability effects HR and how such environmental changes can be used to control the player's experience.

Although a controlled laboratory group was not used in this experiment, this study could greatly benefit from it. In the future when the pandemic ceases, we would like perform this experiment with a control group and with greater number of participants. It would help us in understanding the observed results even further. Current results show promising avenues for future research and has potential for creating better gaming experience for players.

VIII. FUNDING

This research did not receive any specific grant but the author, Sumaira Erum Zaib, receives government scholarship (MEXT) at Tokyo Institute of Technology for doctoral studies.

REFERENCES

- [1] H. Alberts, F. Schneider, and C. Martijn. Dealing efficiently with emotions: Acceptance-based coping with negative emotions requires fewer resources than suppression. *Cognition emotion*, 26:863–70, 01 2012. doi: 10.1080/02699931.2011.625402
- [2] E. Andrade and J. Cohen. On the consumption of negative feelings. *Journal of Consumer Research*, 34:283–300, 02 2007. doi: 10.1086/519498
- [3] H. Araki, T. Ikeda, T. Ozawa, K. Kawahara, and Y. Kawai. Development of a horror game that route branches by the player's pulse rate. In *Proc. Intelligent User Interfaces Companion, IUI '18 Companion*. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3180308.3180322
- [4] K. Bantinaki. The paradox of horror: Fear as a positive emotion. *Journal of Aesthetics and Art Criticism*, 70, 09 2012. doi: 10.1111/j.1540-6245.2012.01530.x
- [5] M. M. Bradley and P. J. Lang. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59, 1994. doi: 10.1016/0005-7916(94)90063-9
- [6] L. Breiman. Machine learning. *Machine Learning*, 45:5–32, 10 2001. doi: 10.1023/A:1010933404324
- [7] M. H. E. M. Browning, K. J. Mimnaugh, C. J. van Riper, H. K. Laurent, and S. M. LaValle. Can simulated nature support mental health? comparing short, single-doses of 360-degree nature videos in virtual reality with the outdoors. *Frontiers in Psychology*, 10, 2020. doi: 10.3389/fpsyg.2019.02667
- [8] G. Calleja. Digital games and escapism. *Games and Culture - Game Cult*, 5:335–353, 09 2010. doi: 10.1177/1555412009360412
- [9] N. Carroll. *The philosophy of horror; or, Paradoxes of the heart*. Routledge, 1990.
- [10] G. Carrus, M. Scopelliti, R. Laforteza, G. Colangelo, F. Ferrini, F. Salbitano, M. Agrimi, L. Portoghesi, P. Semenzato, and G. Sanesi. Go greener, feel better? the positive effects of biodiversity on the well-being of individuals visiting urban and peri-urban green areas. *Landscape and Urban Planning*, 134:221–228, 2015. doi: 10.1016/j.landurbplan.2014.10.022

- [11] E. Champion and A. Dekker. Please biofeed the zombies: Enhancing the gameplay and display of a horror game using biofeedback. In *DIGRA Situated Play*, 2007. doi: 10.25917/5d1443e8af4a0
- [12] E. Chang, H.-T. Kim, and B. Yoo. Virtual reality sickness: A review of causes and measurements. *International Journal of Human-Computer Interaction*, 36:1–25, 07 2020. doi: 10.1080/10447318.2020.1778351
- [13] K.-H. Choi, J. Kim, O. S. Kwon, M. J. Kim, Y. H. Ryu, and J.-E. Park. Is heart rate variability (hrv) an adequate tool for evaluating human emotions? – a focus on the use of the international affective picture system (iaps). *Psychiatry Research*, 251:192–196, 2017. doi: 10.1016/j.psychres.2017.02.025
- [14] M. Clasen, J. Kjeldgaard-Christiansen, and J. Johnson. Horror, personality, and threat simulation: A survey on the psychology of scary media. *Evolutionary Behavioral Sciences*, 14, 11 2018. doi: 10.1037/ebs0000152
- [15] C. Dormann. Affective experiences in the home: measuring emotion. In *Home oriented informatics and telematics*, 2003.
- [16] A. Fischer, P. Rodriguez Mosquera, A. van Vianen, and A. Manstead. Gender and culture differences in emotion. *Emotion (Washington, D.C.)*, 4:87–94, 04 2004. doi: 10.1037/1528-3542.4.1.87
- [17] J. H. Friedman. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29:1189–1232, 2000.
- [18] M. Friedman. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200):675–701, 1937. doi: 10.1080/01621459.1937.10503522
- [19] T. Garner. Identifying habitual statistical features of eeg in response to fear-related stimuli in an audio-only computer video game. 09 2013. doi: 10.1145/2544114.2544129
- [20] T. Garner and M. Grimshaw-Aagaard. A climate of fear: Considerations for designing a virtual acoustic ecology of fear. *AM '11, Audio Mostly 2011, 6th Conference on Interaction with Sound, Coimbra, Portugal, September 7-9, 2011*, pp. 31–38, 09 2011. doi: 10.1145/2095667.2095672
- [21] T. Garner, M. Grimshaw-Aagaard, and D. Nabi. A preliminary experiment to assess the fear value of preselected sound parameters in a survival horror game. *Games Computing and Creative Technologies: Conference Papers (Peer-Reviewed)*, 09 2010. doi: 10.1145/1859799.1859809
- [22] S. Graja, P. Lopes, and G. Chanel. Impact of visual and sound orchestration on physiological arousal and tension in a horror game. *IEEE Transactions on Games*, 13(3):287–299, 2021. doi: 10.1109/TG.2020.3006053
- [23] M. Grimshaw-Aagaard. The audio uncanny valley: Sound, fear and the horror game. *Games Computing and Creative Technologies: Conference Papers (Peer-Reviewed)*, 01 2009.
- [24] M. Hedberg and T. Grip. *Amnesia: The dark decent*, 2010.
- [25] E. Hudlicka. Affective game engines: Motivation and requirements. In *Proc. Foundations of Digital Games*, pp. 299–306, 01 2009. doi: 10.1145/1536513.1536565
- [26] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9):641–661, 2008. doi: 10.1016/j.ijhcs.2008.04.004
- [27] A. Kring and A. Gordon. Sex differences in emotion: Expression, experience, and physiology. *Journal of personality and social psychology*, 74:686–703, 03 1998. doi: 10.1037/0022-3514.74.3.686
- [28] F. Lantz. Exploring the impact of familiarity on the emotional response to acousmatic sound effects in horror film, 2021. doi: 10.13140/RG.2.2.10759.34728
- [29] J.-H. T. Lin. Fear in virtual reality (vr): Fear elements, coping reactions, immediate and next-day fright responses toward a survival horror zombie virtual reality game. *Comput. Hum. Behav.*, 72:350–361, 2017.
- [30] K. E. Madsen. The differential effects of agency on fear induction using a horror-themed video game. *Computers in Human Behavior*, 56:142–146, 2016. doi: 10.1016/j.chb.2015.11.041
- [31] G. K. Mark, D. Alan, and A. Jen. Affective videogames and modes of affective gaming: Assist me, challenge me, emote me. In *DiGRA Changing Views: Worlds in Play*, vol. 3, 2005.
- [32] G. Matan. *Analysis of the Horror Genre and its Implementation in Video Games*. PhD thesis, TH Köln University of Applied Sciences, 06 2019. doi: 10.13140/RG.2.2.10759.34728
- [33] Y. Ng, C. Khong, and H. Thwaites. A review of affective design towards video games. *Procedia - Social and Behavioral Sciences*, 51:687–691, 2012. doi: 10.1016/j.sbspro.2012.08.225
- [34] S. Palma, L. Ripamonti, N. Borghese, D. Maggiorini, and D. Gadia. Player behaviour metrics for adjusting content in vr games: the case of fear. pp. 1–6, 07 2021. doi: 10.1145/3464385.3464705
- [35] B. Perron. *Silent Hill: The Terror Engine*. University of Michigan Press, 2012.
- [36] G. Reznick and L. Fessenden. *Until dawn*, 2015.
- [37] C. Roth, I. Vermeulen, P. Vorderer, and C. Klimmt. Exploring replay value: Shifts and continuities in user experiences between first and second exposure to an interactive story. *Cyberpsychology, behavior and social networking*, 15:378–81, 07 2012. doi: 10.1089/cyber.2011.0437
- [38] G. Roux-Girard. Listening to fear: A study of sound in horror computer games. pp. 192–212, 01 2010. doi: 10.4018/978-1-61692-828-5.ch010
- [39] O. Sagi and L. Rokach. Ensemble learning: A survey. *WIREs Data Mining and Knowledge Discovery*, 8(4):e1249, 2018. doi: 10.1002/widm.1249
- [40] F. Schiltz, C. Masci, T. Agasisti, and D. Horn. Using regression tree ensembles to model interaction effects: a graphical approach. *Applied Economics*, 50(58):6341–6354, 2018. doi: 10.1080/00036846.2018.1489520
- [41] C. Schubert, M. Lambertz, R. Nelesen, W. Bardwell, J.-B. Choi, and J. Dimsdale. Effects of stress on heart rate complexity—a comparison between short-term and chronic stress. *Biological psychology*, 80:325–32, 12 2008. doi: 10.1016/j.biopsycho.2008.11.005
- [42] F. Shaffer and J. P. Ginsberg. An overview of heart rate variability metrics and norms. *Frontiers in Public Health*, 5:258, 2017. doi: 10.3389/fpubh.2017.00258
- [43] N. Simon. *Patterns of Obscurity : Gothic Setting and Light in Resident Evil 4 and Silent Hill 2*. McFarland, 2009.
- [44] M. Slater. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364:3549 – 3557, 2009.
- [45] J. Steuer. Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, 42(4):73–93, 1992. doi: 10.1111/j.1460-2466.1992.tb00812.x
- [46] A. Stuart. Rank correlation methods. by m. g. kendall, 2nd edition. *British Journal of Statistical Psychology*, 9(1):68–68, 1956. doi: 10.1111/j.2044-8317.1956.tb00172.x
- [47] C.-T. Sun, H. Lin, and H.-Y. Lu. Using physiological response data to examine horror video game enjoyment. In *DIGRA Game, Play and the Emerging Ludo-Mix*, 2019.
- [48] J. Takatalo, J. Häkkinen, J. Kaistinen, and G. Nyman. User experience in digital games: Differences between laboratory and home. *Simulation & Gaming*, 42(5):656–673, 2011. doi: 10.1177/1046878110378353
- [49] R. Ueoka and K. Ishigaki. Development of the horror emotion amplification system by means of biofeedback method. In S. Yamamoto, ed., *Human Interface and the Management of Information. Information and Knowledge in Context*, pp. 657–665. Springer International Publishing, Cham, 2015.
- [50] R. Usher, P. Robertson, and R. Sloan. Physical responses (arousal) to audio in games. *The Computer Games Journal*, 2:5–13, 08 2013. doi: 10.1007/BF03392340
- [51] V. Vachirataporn, R. Legaspi, K. Moriyama, K.-i. Fukui, and M. Numao. An analysis of player affect transitions in survival horror games. *Journal on Multimodal User Interfaces*, 9, 12 2014. doi: 10.1007/s12193-014-0153-4
- [52] V. Vachirataporn, R. Legaspi, K. Moriyama, and M. Numao. Towards the design of affective survival horror games: An investigation on player affect. pp. 576–581, 09 2013. doi: 10.1109/ACII.2013.101
- [53] J.-N. Voigt-Antons, R. Spang, T. Kojić, L. Meier, M. Vergari, and S. Möller. Don't worry be happy - using virtual environments to induce emotional states measured by subjective scales and heart rate parameters. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 679–686, 2021. doi: 10.1109/VR50410.2021.00094
- [54] F. Wilcoxon. Individual comparisons by ranking methods. *biometrics bulletin*, 1 (6), 80-83, 1945.
- [55] K. A. Winter and N. A. Kuiper. Individual differences in the experience of emotions. *Clinical Psychology Review*, 17(7):791–821, 1997. doi: 10.1016/S0272-7358(97)00057-3
- [56] Y. Wu, R. Gu, Q. Yang, and Y.-j. Luo. How do amusement, anger and fear influence heart rate and heart rate variability? *Frontiers in Neuroscience*, 13:1131, 2019. doi: 10.3389/fnins.2019.01131