Tuning Stressful Experience in Virtual Reality Games

Susanna Brambilla*
Department of Computer Science,
Università di Milano
Milan, Italy
susanna.brambilla@unimi.it

Giuseppe Boccignone*
Department of Computer Science,
Università di Milano
Milan, Italy
giuseppe.boccignone@unimi.it

N. Alberto Borghese*
Department of Computer Science,
Università di Milano
Milan, Italy
alberto.borghese@unimi.it

Daniele Croci*
Department of Computer Science,
Università di Milano
Milan, Italy
daniele.croci2@studenti.unimi.it

Laura A. Ripamonti*
Department of Computer Science,
Università di Milano
Milan, Italy
ripamonti@di.unimi.it

ABSTRACT

Dynamic game balancing using players' affective state is a promising approach for creating immersive and engaging video games. We present here the results of StrEx, a system which modulates the stress level induced by a video game in a completely unobtrusive way. It has been tested on an *ad-hoc* developed Virtual Reality horror-survival game. The system collects motion behavioral data from Head-Mounted Display and its controllers and updates a dynamical model of the player stress level, that takes into account the interior affective state dynamics. Such model is used to guide the transitions of a Finite State Machine that controls the stress level induced by the game, to maintain an appropriate level of stress. The effectiveness of the system has been evaluated through an experimental study in which participants played the game with and without StrEx. Preliminary results show that the StrEx plugin can increase players' perceived competence and decrease their tension and frustration, thus leading to an increased engagement in the game.

CCS CONCEPTS

• **Human-centered computing** → *Interaction design*; **Interaction techniques**; **Virtual reality**; • **Computing methodologies** → *Artificial intelligence*; Machine learning approaches.

KEYWORDS

video games, virtual reality, technical game design, adaptation, stress, affective computing, behavioural data, dynamical models.

ACM Reference Format:
Susanna Brambilla, Giuseppe Boccignone, N. Alberto Borghese, Daniele Croci, and Laura A. Ripamonti. 2023. Tuning Stressful Experience in Virtual Reality Games. In *15th Biannual Conference of the Italian SIGCHI Chapter (CHItaly 2023), September 20–22, 2023, Torino, Italy*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3605390.3605412

*All authors contributed equally to this research.

CHItaly 2023, September 20–22, 2023, Torino, Italy

ACM ISBN 979-8-4007-0806-0/23/09…$15.00
https://doi.org/10.1145/3605390.3605412

1 INTRODUCTION

Today's research in technical video game design is increasingly focusing on making games adaptive, so that the gaming experience can become as enjoyable as possible for anyone, regardless of his/her skills and experience [3]. One concept that captures the essence of a perfect gaming experience is Csikszentmihalyi's idea of "being in the Flow" [16]. Flow is a mental state in which players are in such a high level of concentration that they are completely absorbed in the activity. This state is achieved when players' abilities and expectations are evenly matched with the challenge posed by the game: a too low challenge will lead players to boredom, while a challenge level that overwhelms the player will induce frustration and anxiety.

Anyway, achieving an adequate game balance is a difficult and time-consuming process: generally, it is iteratively refined by tweaking by hand specific game behaviors and parameters on the basis of the feedback received during playtesting sessions [37]. This limit has been partially overcame by the Dynamic Difficulty Adaptation (DDA) approaches [43, 45], which aim at optimizing the experience for the player by adapting the game content on the fly, mainly using techniques borrowed from Artificial Intelligence (AI) to modulate game parameters in order to avoid the player getting out of the Flow channel [20, 32]. Given the existence of many different types of players [8], adaptation relying only upon in-game performance might achieve a limited success. Each player brings forward his/her own goals, preferences, personality, skills and prior emotional/cognitive competence (in a nutshell, a player internal model [31]) while playing. In this respect, it has been argued that the use of data regarding the emotional state of each player has a key role in achieving an effective game balance [41, 43]. Unfortunately, this approach is too often vaguely applied, due to the difficulty of clearly tracking and distinguishing between emotions, affect and stress (the latter being a major concern for gaming - markedly, with reference to the flow [30]). Considering the subtleties of the subject, we start establishing a clear definition of the core terms and concepts exploited in the remainder of the paper.

A player, given his/her affective internal model, might perceive a situation as stressful whenever a discrepancy occurs between the player's expectations and the situation outcomes, and the resources (actions) available to minimise this discrepancy are inadequate [23, 24, 31]. Henceforth, action inefficacy often triggers primary stress responses involving either the unfolding of negative emotions (e.g., "anger") or physiological reactions [23, 24]. Emotions, typically labelled by terms such as "fear", "anger", and so forth, can be regarded as constructed categories by means of which we conceptualise, in a given context, the complex entanglement of a number of components, such as intero/exteroception, core affect, instrumental actions, physiological/behavioural changes, subjective conscious experience and categorisation [4, 6, 7, 11]. Of particular relevance here is the core affect [1, 5, 6, 39], a psychological construct best defined as a mental state of pleasure or displeasure (valence), with some degree of activation (arousal) (see, e.g., Russell's circumplex model [38]). The core affect emerges from the intertwinement of exteroceptive and interoceptive sensations and it is grounded in the *internal milieu*: the integrated sensory representation of the physiological state of the body [1, 5, 6, 39]. As such by no means emotions can be reduced to core affective states or vice versa [7]. Such distinction should always be kept in mind, beyond its theoretical relevance, because it bears profound consequences in the development of emotion-based Human-Machine Interaction. For instance, we cannot measure changes in emotions when the measurement tool only allows inferences about core affect [35]. For the same reason, physiological measurements should be handled with care; indeed bodily physiological changes are multiply determined by both sympathetic and parasympathetic autonomic changes that are sensitive to many psychological factors other than just affect or emotion - for instance, attention or mental effort - or even crudely caused by changes in the physical environment [35]. Under such circumstances, the gaming emotional experience unfolds as the dynamics of a complex interaction system [36]: by designing a game we set a cognitive/emotional context inside which each gamer, at any moment, is expected to face challenges with his/her own internally available (predictive) model armed with affordable actions. The greater or minor ability to cope with such challenges impacts on the player's stress response together with the general affect state changes and specifically perceived emotions. These, in turn, impinge on his/her behavioural and physiological (re)actions.

Brambilla et al. have recently shown [13] how behavioural data can be exploited as observable variables to infer a model of the latent dynamics of the player's stress level. The focus of this paper is to present the results of utilizing Stressful Experience (StrEx), a system for dynamically adapting the video game content using the stress level provided by [13]. Based on their work, we have designed and developed an *ad-hoc* first-person survival horror game using *Unity 3D*. The choice of the game genre has been inspired by [42], which underlines its effectiveness in eliciting strong emotional responses. We have also conducted a study with a group of testers who played the video game while collecting data. The preliminary results have allowed us to assess the validity of StrEx.

The remainder of this work is organized as follows. In Section 2, the state-of-the-art in the related research is outlined. Section 3 briefly describes our approach and the prototype video game we have designed and developed for testing purposes. Section 4 presents the StrEx plugin in details. Section 5 summarizes our experimental procedure and reports the results achieved so far. Finally, in Section 6 some conclusions are drawn.

2 STATE-OF-THE-ART

Few approaches have been proposed to evaluate the affective state of players while playing [2]. Frommel et al. [17] developed *Space-Jump*, a 2D scrolling down platform web-based game, whose difficulty level is adjusted from players' performance. The game included several dialogues with Non-Player Characters (NPCs) aimed at unobtrusively assessing how much the player was frustrated and/or bored. Results demonstrated the effectiveness of the adaptation heuristics to keep the player into the flow, but also highlighted that this approach is perceived as too invasive, thus producing a negative effect on the overall player experience.

Nogueira et al.[28] proposed an approach to identify valence and arousal from physiological data. To this aim they gathered such data during the exposition to different stimuli: relaxing music, film clips, and scaring images. Participants were then asked to rate their perceived arousal/valence associated with the stimuli. During the gaming sessions, physiological data were acquired and mapped to valence/arousal values; the latter were used to adapt at run-time the generation of levels, assets, events, enemies' behaviour, character attributes, sound and visual effects, light sources, and item placement. Results with and without game adaptation were evaluated via the Game Experience Questionnaire (GEQ)[34, 41]. Players appreciation raised with adaptation; this was especially true for experienced players that liked the horror genre. Nevertheless, we remark that in this approach, mapping is built inside a non-interactive context (passively watching or hearing stimuli), while affective response may vary significantly in the context of an interactive activity, such as a video game.

Several proposal to adjust game difficulty from emotion-related data have been developed also by practitioners. One notable example is the so-called AI-Director developed by Valve for its *Left 4 Dead* (Valve, 2008), a cooperative First-Person Shooter (FPS) game. The AI Director manages enemies position and quantity, availability of resources, music and visual effects. The goal is to maximize players' amusement by alternating between frenzy and more relaxed gameplay phases, thus implementing an effective "dramatic curve" [40]. To define a proxy of the emotional state of the player, the AI Director calculates a "stress value" based on a heuristics on different in-game events (namely: damage taken by players, number of enemies killed, and so forth) and adjusts the game content accordingly[1]. As a result, at each playthrough the players are faced with a slightly different situation, depending on how well they are performing, thus stimulating them to play again and again the same levels - and thus stretching the product life-cycle. Although a pioneering approach for its time, the AI Director suffered from several limitations, mainly deriving from how the proxy had been designed and from the declared goal of the project. Actually, Valve developed its systems to improve the "replyability" of the game, and not to adapt the content to the players' real affective state.

[1] But see: https://cdn.cloudflare.steamstatic.com/apps/valve/2009/GDC2009_ReplayableCooperativeGameDesign_Left4Dead.pdf (Mike Booth, Game Developers Conference, 2009).

A different approach has been pioneered in *Hiramon* [18]: the idea was to detect players' emotions directly from their interaction with the gaming device (to avoid asking players to wear any specific sensor). *Hiramon* was a serious game aimed at teaching how to write Japanese hiragana characters. Players were first shown how the characters are written and then asked to replicate them. Then, during the game, "enemies" challenged players to correctly write characters randomly selected. To overcome the enemy, players must write at least three characters out of five without errors. During these phases, the game collected data about both the player performance and his/her interaction with the device (a tablet), in terms of pressure and strokes. Moreover, after each fight, players were asked to self-report their affective state in terms of valence/arousal/dominance. Note that, albeit they declare to aim at emotions, the authors are instead assessing affect (cfr. Section 1) classically defined in terms of valence, arousal and dominance [38]. Data are then used to train a Random Forest-based classifier to predict discretized levels of valence, arousal, and dominance that achieved an accuracy up to 74%. The difficulty of scoring a relevant classification rate may be due to the fact that the approach neglects the important role of affective state dynamics [9, 21] and focuses on a simplified input-output, stateless, approach.

In this paper, we propose a work that extends a previous study by testing and evaluating the Strex plugin. We push the approach of [18] one step further, using the plugin to detect perceived stress in dynamic game balancing exploiting sensor data from gaming devices. By leveraging the richer interaction data and higher immersion provided by a VR game, we propose a non-obtrusive approach to estimate the player's stress level. Moreover, adopting a dynamical model to estimate the stress level of the player we aim at overcoming some of the limitations of the here above approaches and achieving a higher accuracy in stress level estimate.

Table 1: Feature list (✓= Feature used to feed the models, corr = correlation, ampl = amplitude, vel = velocity, acc = acceleration, ang = angular, press = pressure, pos = position, # = number). Source: [13].

TYPE	DATA	FEATURE				
		μ	σ	*min*	*max*	$\mu\#$
head	vel	✓	✓	✓	✓	
	ang vel	✓	✓	✓	✓	
	acc	✓	✓	✓	✓	
	ang acc	✓	✓	✓	✓	
left hand	vel	✓	✓	✓	✓	
	ang vel	✓		✓		
	acc	✓	✓	✓	✓	
	ang acc	✓	✓	✓	✓	
	grip press	✓	✓	✓	✓	
	trigger press	✓	✓		✓	
	thumbstick pos x	✓	✓	✓	✓	
	thumbstick pos y	✓	✓	✓	✓	
	grip pressed					✓
	trigger pressed					✓
right hand	velocity	✓	✓	✓	✓	
	ang vel	✓	✓	✓	✓	
	acc	✓	✓	✓	✓	
	ang acc	✓	✓	✓	✓	
	grip press	✓		✓	✓	
	trigger press	✓	✓	✓	✓	
	thumbstick pos x			✓		
	thumbstick pos y					
	grip pressed					✓
	trigger pressed					✓

3 DESIGN AND EVALUATION OF STREX

StrEx is conceived as a plugin for games that capitalizes on the model developed in [13] (see Fig.1, box *a*) to estimate the stress level of the player. The predicted stress level is then used to adapt the difficulty level of a specifically developed game, as shown in box *b* of Fig.1.

To elicit stronger and more easily identifiable emotional response, we have designed and developed a prototype horror-survival game, titled *StrEx Space*. This game is specifically designed for Virtual Reality (VR) to enhance player engagement and sense of presence in simulated environments [25, 29]. Moreover, we chosed Meta Quest 2[2] as VR device, since it is able to track and acquire a rich repertoire of motion data (cf. Tab.1).

The model incorporated into StrEx (Fig.1, box *a*), was trained using players' motion behavioral data acquired by the sensors of Meta Quest 2 while testers were playing a simplified version of *StrEx Space* (with no adaptation) and players self-annotated their perceived stress level through an *ad-hoc* version of *DANTE*[3] (Dimensional ANnotation Tool for Emotions) [10]. The tool includes sliders for continuous annotation of valence and arousal on a scale ranging from -1 to 1, with a step size of 0.001. For this particular study, arousal and valence sliders were not used, and DANTE was expanded to include a stress slider. This addition enabled participants to annotate their perceived stress levels using a continuous extension of a Likert scale, represented as a color bar. The scale ranges from low stress (-1) to high stress (1).

Motion behavioral data were collected in real-time, sampling them at 64Hz and extracting features from sliding windows of 6s, with an overlapping of 5s. Features were selected based on their correlations with the stress level (see Tab.1), calculated using Pearson correlation coefficient. The features vector obtained was then fed into a Hidden Markov Model (HMM) which returned a binary value indicating the presence/absence of stress with an accuracy of 84.4%.

3.1 StrEx Space Gameplay Overview

StrEx has been developed taking inspiration from [33], with the aim of providing different stress levels and a convincing game play.

To escape an abandoned space station invaded by monstrous enemies, players must find an escape pod and solve several puzzles to reactivate it. The game map is procedurally generated at the beginning of each session: it has a labyrinthine structure composed of rooms and corridors. Rooms may contain either enemies, collectable items (e.g., ammunition, medical kits) or both. These are spawned randomly on the fly, increasing/decreasing their spawning

[2] https://store.facebook.com/it/quest/products/quest-2/
[3] https://github.com/phuselab/DANTE

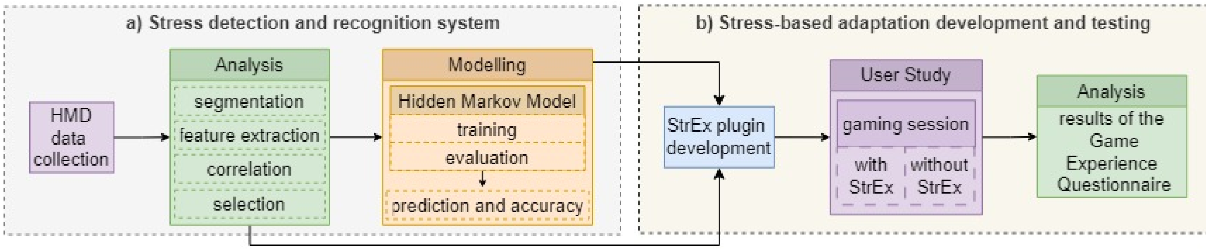

Figure 1: Box a) summarizes the HMM-based method proposed in [13], which is exploited here to evaluate the latent stress level of the player: it is able to predict the presence/absence of stress with an accuracy of 84.4%. Box b) shows the outline of the study reported here: the same type of data acquired for training the model is used here to adapt the content of *StrEx Space* game through the trained stress model. Results on game experience with and without adaptation are analyzed.

frequency with the current player's progress and stress level. In order to survive, players have to preserve their health and reserve of oxygen. Health decreases (leading eventually to death) when a player is injured by enemies, but can be partially restored using medical kits. Oxygen decreases constantly, but its consumption rate is affected by the stress level of the player: higher stress implies higher consumption. Oxygen reserve can be replenished using oxygen cylinders found in the game, anyway, when the oxygen reserve drops to zero, health starts to decline very quickly.

3.2 Design of Stressors

Drawing on Lebois et al. [24], we have introduced in StrEx Space several *ad-hoc* stressors, using distinctive elements of the specific game-genre: sense of isolation, presence of monstrous enemies (see, e.g., Fig.2), threatening sound and music, lack of visibility, lack of oxygen, and scarcity of resources. The rationale for adaptation is that the larger the number of stressors the higher is the stress level. In fact, as reported in Section 1, a situation is perceived as stressful whenever a discrepancy occurs between expectations of an individual and the specific situation he or she is going through [23], [19, 31].

Figure 2: Example of stressors in the game

4 STREX PLUGIN STRUCTURE

StrEx includes two modules: the first one (*Stress Model*) implements the stress detection model, whilst the other (*Content Adapter*) uses AI-based techniques to handle game content according to the stress module outcome.

The first module gathers real-time data from the Meta Quest 2, and extract the features to feed a HMM that returns a binary value (stress/no stress). This value is passed on to the Content Adapter module, which implements a Finite State Machine (FSM) inspired by the *Left 4 Dead* AI-Director[12]. Four states are considered: *Build Up*, *Sustain Peak*, *Peak Fade* and *Relax* (Fig. 3).

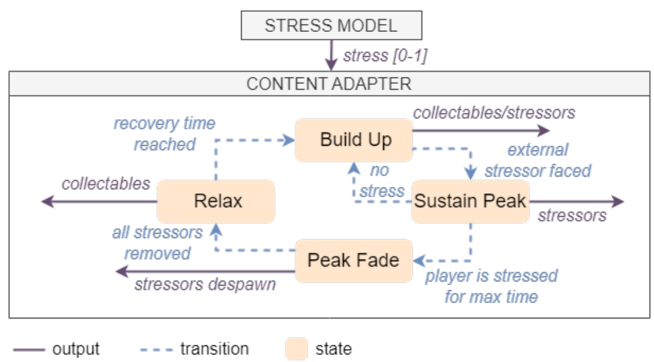

Figure 3: Scheme of the plugin with an emphasis on the FSM of the Content Adapter module.

The transition between states is determined by the following inputs: *stress level provided by the model, stressful events faced by the player* and *time spent inside a state.*

In each state, a set of stressors are spawned or removed. These can be classified as *internal* or *external* according to their distance from the player. Internal stressors are quick disturbing events (e.g., scary sounds, lights flickering) that are spawned inside the current room and are not directly related to advancing inside the game but they do contribute to the build up of stress. External stressors are obstacles (e.g., enemies) that must be faced by the player to advance in the game and are spawned in adjacent rooms.

The choice of the stressor to be generated is performed randomly, to prevent boring the player with the repetition of the same event.

The starting state is Build Up state and we here describe what happens when the FSM enters in each state.

4.1 FSM States

The plugin starts in the *Build Up* state, where stressors and collectible items are generated in the player's current area. If the player face to confront an external stressor, the FSM switches to the *Sustain Peak* state, whichever is the outcome of the fight.

In the *Sustain Peak* state, a timer starts to measure the time spent in this state. At each time step, the timer and the stress model's output are evaluated. If the player is not stressed, the FSM returns to the *Build Up* state. On the contrary, if the player is stressed, the FSM remains in the *Sustain Peak* state until the maximum allowed time (typically between 3 and 10 seconds), based on the desired stress level, is reached.

Afterwards, the FSM moves to the *Fade* state, where no further stressors are spawned, and any pending stressors are removed. The FSM then switches to the *Relax* state.

In the *Relax* state, the objective is to calm down the player. No stressors are generated, and the number of useful items that can be found is increased. Once the player has traveled a distance equivalent to two rooms or after a predefined recovery time, the FSM reverts to the *Build Up* state and the cycle repeats. The recovery time is based on the Left 4 Dead AI Director, typically ranging from 30 to 45 seconds.

4.2 Collectable item Generation Process

The presence of useful items aims to alleviate the player's stress and empower them. Indeed, item generation takes place mainly in the *Relax* state. The collectable items generation process follows a Goal-Oriented Behaviors approach [26]. In this case the *goals* are the items to be generated, the *actions* are their actual generation and the *insistence* represents the level of importance of each item, in the current game situation. During each iteration, a search is performed in adjacent rooms to find points where resources can be spawned. The item with the highest insistence value is selected with a probability of being spawned directly proportional to its insistence (up to 60%). Additionally, a 10% probability of spawning the item is added if the FSM is in the *Relax* state.

5 RESULTS

To evaluate StrEx effectiveness, we have set up a testing session organized as shown in Fig.4. Main goal was to determine to which extent the adaptation technique described here was effective in keeping the player into the flow.

We recruited 13 testers - mainly among Ms students in Computer Science: 12 males and 1 female, in the age range 18 - 29. All of them - except one - play regularly video games (on average 10 hours a week). Their favourite game genre is role-playing, followed by action and adventure games. All the testers except 3 stated to appreciate horror survival-themed games; 3 participants had never experienced VR games before; none suffered from anxiety disorder. All participants volunteered and received no payment.

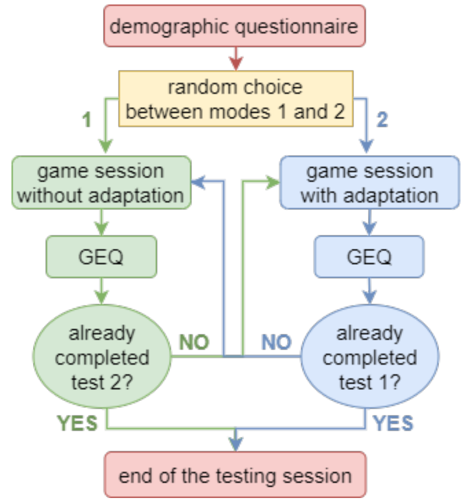

Figure 4: Flow chart of a testing session

5.1 Testing Procedure

Testers were asked to play StrEx Space two times, as shown in Fig.4. One game session was based on a predetermined map (in which events take place without taking into account the stress level of the player) while the other exploited the adaptation techniques we have implemented. To reduce bias, testers were assigned to play the fixed or the adapted version of the game in random order and they were not aware which version they were playing. The gaming session lasted approximately 30-40 minutes for each mode.

At the end of each play session, each participant was asked to complete the core module of the Game Experience Questionnaire (GEQ) [34, 41], which is aimed at measuring the emotional state and thoughts of players while playing and after playing a game. GEQ is composed of several modules, which can be used collectively or not:

- *core* (GEQ): probes multiple components of players' experience while gaming;
- *post-game* (PGQ): investigates gamers' experience and possible other after effects after the gaming session;
- *social presence* (SPGQ): evaluates gamers' experience and involvement with co-player(s) in multiplayer games;
- *in-game GEQ* (iGEQ): probes in-game experience multiple times during the same gaming session.

For our purposes, the *core module* was enough, since we were not interested in post-game effects, our game had no social interaction and we wanted to avoid disrupting the immersivity of the experience with an in-game questionnaire, which may have resulted especially invasive and cumbersome when playing a VR game.

5.2 Data Analysis

The core GEQ explores the players experience along 7 dimensions (called "components"), using 42 questions whose answer is provided on 5-points Likert scale. Each component is probed by a specific subset of questions whose average score represents the score of the component (see [34]). The components analysed by GEQ are:

- *sensory and imaginative immersion*: richness of the experience from the aesthetic, storytelling, sensory and imaginative points of view;
- *flow*: concentration and absorption of the player in the game;
- *competence*: players' perception of their own skills and abilities;
- *positive affect*: how much players like the game and how much fun they are having;
- *negative affect*: players' negative experiences with the game;
- *tension/annoyance*: pressure and frustration felt by the players;
- *challenge*: level of challenge perceived by the players.

We have not considered here the sensory and imaginative component, since it is related primary to the game storytelling and to the aesthetic aspects of the game. StrEx space has just a short premise (no real story) and the aesthetic is not relevant for the present work. Moreover, both remain unaltered between the two versions of the level played by the tester, therefore any comparison would be meaningless.

Distribution of the average score of the different components of the core GEQ are reported in violin plots of Fig. 5.

Competence median increases in mode 2, that implies that players felt more successful and at ease with the stressors dynamics used when adaptation is in place.

The variance of the *flow* component shrinks, but the median is similar in the two game modes. This result, apparently in contrast with the trend of competence, could be due to the misinterpretation of question 25 ("I lost track of time"), whose score affects the total score of the flow component. As can be seen in Fig. 6), answers to this specific questions are not aligned to those of the remaining questions in this sub-group. Our interpretation is that, in the peculiar context of our horror games, asking to testers whether they have lost track of time could be misleading, since it could be easily interpreted as unwilling to run out of time to escape from the station due to oxygen decrease. If we take out the answer to this specific question, the mean of the flow increase when adaptation is considered.

Also both *negative* and *positive affect* increase in mode 2. This apparently contradictory result supports the view that positive and negative emotions can co-exist inside us (cf. Larsen et al. [22]). *Negative Affect* generally increases when adaptation is in place. It has a low value, mainly in the Relax state, and it increase for the Build Up and Sustained Peak states where a larger presence of stressors take place. We remark that horror games could generate negative affective states, even when players are enjoining the game, and, at the same time, positive affective states could be generated by horror elements, especially in the video game context [27].

Support to the positive effect of adaptation is also provided by *tension/annoyance* and *challenge* whose median decreases in mode 2. Both components are related to flow, and their decrease shows that the player felt, on average, less bored and stressed, and also more skillful while gaming.

Since the number of participants in the experiment was limited, a plain statistical analysis between groups of players with different characteristics is beyond reach. Hence, we have examined the answers of each participant with the aim of gaining a better understanding of the experience they had[4]. To this aim, participants were divided into groups based on: their experience with VR games, their appreciation for horror games, and the average number of hours played per week. In general, all participant groups felt more *competent* playing the game with adaptation. As said, competence is related to the self-perceived skill level, hence it could be an effective indicator of the fact that players were involved in the game.

No large difference between groups was observed for the different components. *Competence*: the participants of all groups felt more *competent* playing the game with adaptation. The same is true for *flow* component, that does not exhibit major changes. It slightly decreases for non-horror lovers and for participants who declared to have a substantial experience with VR, and slightly increases for participants who had no previous experience with VR. Players with no experience with VR, players who play less hours per week, and players who do not like horror games are the groups which evidenced the greater improvements on *Tension/Annoyance* component, which anyway decreases across all groups of participants. This component is strongly related to the concept of flow, hence this result is positive.

6 DISCUSSION AND CONCLUSION

Preliminary results reported here suggest that the dynamic adjustment of game events based on players stress level leads to an improvement in their perceived competence. This, along with the slight decrease in tension/annoyance and challenge, indicates that the game overall becomes less difficult and frustrating for players and therefore more enjoyable. This with adaptation kept at a minimum: it is used in two ways. First, it determines how long the Sustain Peak state should be kept as the player is brought back to Build Up state until he/she becomes stressed, and this contributes in providing a challenge level adequate to induce stress in the player. Secondly, it determines the density of collectable items in the Relax State, that ultimately may increase the confidence of the player.

Quite interesting, both negative and positive affect is elicited. This should be regarded, to some extent, as not surprising as the game has horror components, that could elicit both positive and negative affective states, that are associated to positive and negative valence of the emotional state. A longstanding viewpoint in psychology is that valence is a one-dimensional concept within the space of core affect [38], with the other dimension being physiological arousal or subjective intensity. However, it has been argued that valence is actually a two-dimensional construct [14, 15], with each axis representing the intensity of negative and positive valence. This approach acknowledges that emotions can be complex categories that may allow for, at the affect level, the coexistence of both positive and negative components, leading to a more nuanced understanding of emotional experiences. Most related to the work here, Larsen et al. [22] suggested that special circumstances could enhance mixed emotions of positive and negative affect. They collected continuous measures of positive and negative affect by participants, showing that disappointing wins and relieving losses elicit mixed emotions simultaneously. In conclusion, the preliminary results obtained so far by experimenting StrEx seems

[4]All data are freely accessible here: git@pong.di.unimi.it:research/stress.

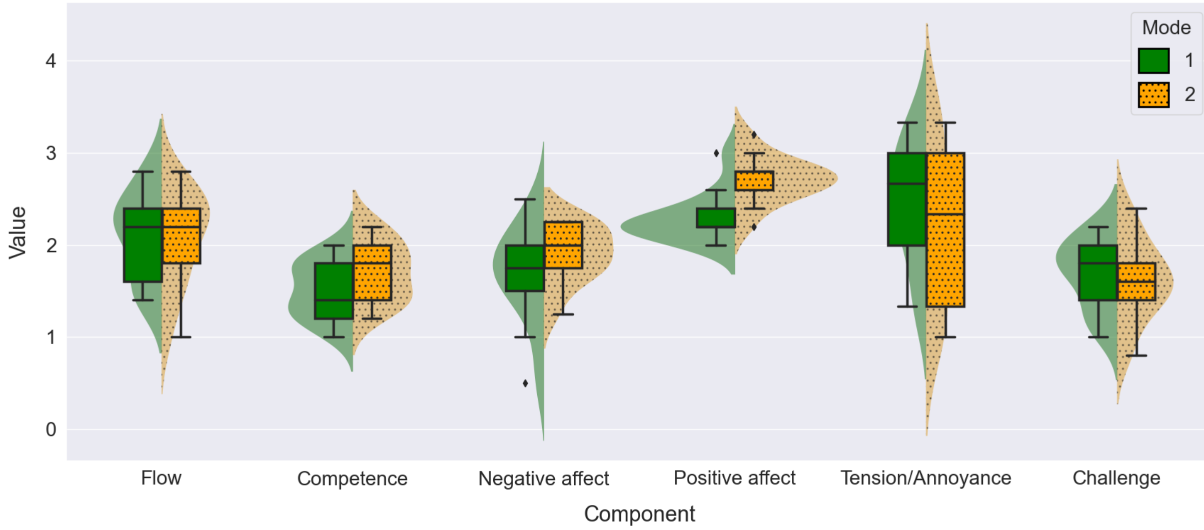

Figure 5: Violin plots, computed from the GEQ data of all participants when adaptation is not used (Mode 1) and when it is used (Mode 2). The violin silhouette visualizes the density estimate of the underlying score empirical distribution; the plot is overlaid with a box plot - displaying 1st, 2nd (median) and 3rd quartiles, maximum and minimum values (whiskers), outliers as dots. Each plot focuses on one specific component of the GEQ.

	FLOW									
	5) I was fully occupied with the game		13) I forgot everything around me		25) I lost track of time		28) I was deeply concentrated in the game		31) I lost connection with the outside world	
Subj	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2
1	2	3	3	4	3	4	3	4	3	3
2	4	4	4	4	4	4	4	4	4	4
3	4	3	3	2	2	1	3	2	3	2
4	4	4	3	3	0	1	3	3	2	2
5	4	4	2	2	3	1	4	3	3	3
6	3	4	3	4	3	2	3	3	3	4
7	4	4	3	3	3	1	4	4	2	4
8	4	4	3	4	3	1	3	3	3	4
9	3	3	4	4	4	3	3	4	3	3
10	3	3	3	3	3	4	3	3	3	3
11	4	4	3	4	1	1	3	3	2	3
12	4	4	4	4	4	4	4	4	4	4
13	3	3	3	3	4	3	4	3	3	3

values of M1 values which didn't changed between modes

Values improved in M2 (1 – 3) Values worsened in M2 (1 – 2)

Figure 6: Results of GEQ questions related to flow component and for each subject (M1 = mode without adaptation, M2 = mode with adaptation).

encouraging; however, further tests and applications in different games genres would be needed to validate its actual effectiveness.

We are currently developing more refined prototype games that can exploit adaptation in several ways, while further developing the stress model to include other measures of stress, such as those that can be obtained by analysing the voice of the player. In that way, we are aiming to cross-validate the data used to classify the stress level of the player, while further refining the classification. In a similar vein, another improvement we are aiming at is to design a continuous stress inference system, able to detect not only the presence of stress, but also to distinguish its intensity and to track its dynamical evolution in time. This would allow us to build an adaptable balancing system based on a reliable real-time assessment of the player's stress state. Last but not least, another relevant open issue is personalization. As clearly shown by individual scores from the core GEQ 5, players with different characteristics or playing styles have different experience-related responses to the game: clustering users into homogeneous groups [44] would allow to offer an *ad-hoc* experience tailored to each group.

ACKNOWLEDGMENTS

This work has been partially supported by EC H2020 ESSENCE project, Grant number 101016112.

REFERENCES

[1] David J Anderson and Ralph Adolphs. 2014. A framework for studying emotions across species. *Cell* 157, 1 (2014), 187–200.
[2] Renan Vinicius Aranha, Cléber Gimenez Corrêa, and Fátima LS Nunes. 2019. Adapting software with affective computing: a systematic review. *IEEE Transactions on Affective Computing* 12, 4 (2019), 883–899.
[3] Sander Bakkes, Chek Tien Tan, and Yusuf Pisan. 2012. Personalised gaming: a motivation and overview of literature. In *Proceedings of the 8th Australasian Conference on Interactive Entertainment: Playing the System*. 1–10.
[4] Lisa Feldman Barrett. 2006. Are emotions natural kinds? *Perspectives on psychological science* 1, 1 (2006), 28–58.
[5] Lisa Feldman Barrett and Eliza Bliss-Moreau. 2009. Affect as a psychological primitive. *Advances in experimental social psychology* 41 (2009), 167–218.
[6] Lisa Feldman Barrett and James A Russell. 2014. *The psychological construction of emotion.* Guilford Publications.
[7] Lisa Feldman Barrett and Ajay B Satpute. 2019. Historical pitfalls and new directions in the neuroscience of emotion. *Neuroscience letters* 693 (2019), 9–18.

[8] Richard Bartle. 1996. Hearts, clubs, diamonds, spades: Players who suit MUDs. *Journal of MUD research* 1, 1 (1996), 19.
[9] Giuseppe Boccignone, Donatello Conte, Vittorio Cuculo, Alessandro D'Amelio, Giuliano Grossi, and Raffaella Lanzarotti. 2018. Deep construction of an affective latent space via multimodal enactment. *IEEE Transactions on Cognitive and Developmental Systems* 10, 4 (2018), 865–880.
[10] Giuseppe Boccignone, Donatello Conte, Vittorio Cuculo, and Raffaella Lanzarotti. 2017. AMHUSE: a multimodal dataset for HUmour SEnsing. In *Proc of the 19th ACM International Conference on Multimodal Interaction*. 438–445.
[11] Michael Boiger, Eva Ceulemans, Jozefien De Leersnyder, Yukiko Uchida, Vinai Norasakkunkit, and Batja Mesquita. 2018. Beyond essentialism: Cultural differences in emotions revisited. *Emotion* 18, 8 (2018), 1142.
[12] Michael Booth. 2009. The ai systems of left 4 dead. In *Artificial Intelligence and Interactive Digital Entertainment Conference at Stanford, 2009*.
[13] Susanna Brambilla, Giuseppe Boccignone, Alberto N. Borghese, and Laura A. Ripamonti. 2022. Between the Buttons: Stress Assessment in Video Games using Players' Behavioural Data. In *Proceedings of the 6th International Conference on Computer-Human Interaction Research and Applications - CHIRA*. INSTICC, SciTePress, 59–69. https://doi.org/10.5220/0011546400003323
[14] Benny B Briesemeister, Lars Kuchinke, and Arthur M Jacobs. 2012. Emotional valence: A bipolar continuum or two independent dimensions? *Sage Open* 2, 4 (2012), 2158244012466558.
[15] John T Cacioppo and Gary G Berntson. 1994. Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological bulletin* 115, 3 (1994), 401.
[16] Mihaly Csikszentmihalyi. 1990. *Flow: The psychology of optimal experience*. Vol. 1990. Harper & Row New York.
[17] Julian Frommel, Fabian Fischbach, Katja Rogers, and Michael Weber. 2018. Emotion-based dynamic difficulty adjustment using parameterized difficulty and self-reports of emotion. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. 163–171.
[18] Julian Frommel, Claudia Schrader, and Michael Weber. 2018. Towards emotion-based adaptive games: Emotion recognition via input and performance features. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. 173–185.
[19] Tory E. Higgins. 1989. Self-discrepancy theory: What patterns of self-beliefs cause people to suffer? *Advances in Experimental Social Psychology* 22 (1989), 93–136.
[20] Raph Koster. 2013. *Theory of fun for game design*. O'Reilly Media, Inc.
[21] Peter Kuppens, Zita Oravecz, and Francis Tuerlinckx. 2010. Feelings change: Accounting for individual differences in the temporal dynamics of affect. *Journal of personality and social psychology* 99, 6 (2010), 1042.
[22] Jeff T Larsen, A Peter McGraw, Barbara A Mellers, and John T Cacioppo. 2004. The agony of victory and thrill of defeat: Mixed emotional reactions to disappointing wins and relieving losses. *Psychological science* 15, 5 (2004), 325–330.
[23] Richard S. Lazarus. 1993. From Psychological Stress to the Emotions: A History of Changing Outlooks Annual Review of Psychology. *Annual Review of Psychology* 44 (1993), 1–22.
[24] Lauren AM Lebois, Christopher Hertzog, George M Slavich, Lisa Feldman Barrett, and Lawrence W Barsalou. 2016. Establishing the situated features associated with perceived stress. *Acta Psychologica* 169 (2016), 119–132.
[25] Javier Marín-Morales, Carmen Llinares, Jaime Guixeres, and Mariano Alcañiz. 2020. Emotion recognition in immersive virtual reality: From statistics to affective computing. *Sensors* 20, 18 (2020), 5163.
[26] Ian Millington and John Funge. 2018. *Artificial intelligence for games*. CRC Press.
[27] Yousif Nash. 2015. *Slay the Dragon: Writing Great Video Games, Robert Denton Bryant and Keith Giglio*. Michael Wiese Productions.
[28] Pedro A Nogueira, Vasco Torres, Rui Rodrigues, Eugénio Oliveira, and Lennart E Nacke. 2016. Vanishing scares: biofeedback modulation of affective player experiences in a procedural horror game. *Journal on Multimodal User Interfaces* 10, 1 (2016), 31–62.
[29] Federica Pallavicini, Ambra Ferrari, Alessandro Pepe, Giacomo Garcea, Andrea Zanacchi, and Fabrizia Mantovani. 2018. Effectiveness of virtual reality survival horror games for the emotional elicitation: Preliminary insights using Resident Evil 7: Biohazard. In *International Conference on Universal Access in Human-Computer Interaction*. Springer, 87–101.
[30] Corinna Peifer, André Schulz, Hartmut Schächinger, Nicola Baumann, and Conny H Antoni. 2014. The relation of flow-experience and physiological arousal under stress—can u shape it? *Journal of Experimental Social Psychology* 53 (2014), 62–69.
[31] Achim Peters, Bruce S McEwen, and Karl Friston. 2017. Uncertainty and stress: Why it causes diseases and how it is mastered by the brain. *Progress in neurobiology* 156 (2017), 164–188.
[32] Michele Pirovano, Renato Mainetti, Gabriel Baud-Bovy, Pier Luca Lanzi, and N Alberto Borghese. 2014. Intelligent game engine for rehabilitation (IGER). *IEEE Transactions on Computational Intelligence and AI in Games* 8, 1 (2014), 43–55.
[33] Michele Pirovano, Elif Surer, Renato Mainetti, Pier Luca Lanzi, and N Alberto Borghese. 2016. Exergaming and rehabilitation: A methodology for the design of effective and safe therapeutic exergames. *Entertainment Computing* 14 (2016), 55–65.
[34] Karolien Poels, Yvonne A.W. de Kort, and Wijnard A. IJsselsteijn. 2007. *D3.3: Game Experience Questionnaire: development of a self-report measure to assess the psychological impact of digital games*. Technische Universiteit Eindhoven.
[35] Karen S Quigley, Kristen A Lindquist, and Lisa Feldman Barrett. 2014. Inducing and measuring emotion and affect: Tips, tricks, and secrets. In *Handbook of research methods in social and personality psychology*, H. T. Reis and C. M. Judd (Eds.). Cambridge University Press, New York, NY, USA, 220–252.
[36] Michael J Richardson, Rick Dale, and Kerry L Marsh. 2014. Complex dynamical systems in social and personality psychology: Theory, modeling, and analysis. In *Handbook of research methods in social and personality psychology*, H. T. Reis and C. M. Judd (Eds.). Cambridge University Press, New York, NY, USA, 253–282.
[37] Andrew Rollings and Ernest Adams. 2003. *On Game Design*. New Riders.
[38] James A Russell. 1980. A circumplex model of affect. *Journal of personality and social psychology* 39, 6 (1980), 1161.
[39] James A. Russell. 2003. Core affect and the psychological construction of emotion. *Psychological review* 110, 1 (2003), 145.
[40] Jesse Schell. 2008. *The Art of Game Design: A book of lenses*. CRC press.
[41] Tim JW Tijs, Dirk Brokken, and Wijnand A IJsselsteijn. 2008. Dynamic game balancing by recognizing affect. In *International Conference on Fun and Games*. Springer, 88–93.
[42] Vanus Vachiratamporn, Roberto Legaspi, Koichi Moriyama, and Masayuki Numao. 2013. Towards the design of affective survival horror games: An investigation on player affect. *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction* (2013), 576–581.
[43] Georgios N Yannakakis and John Hallam. 2009. Real-time game adaptation for optimizing player satisfaction. *IEEE Transactions on Computational Intelligence and AI in Games* 1, 2 (2009), 121–133.
[44] Georgios N Yannakakis and Julian Togelius. 2018. *Artificial intelligence and games*. Vol. 2. Springer.
[45] Mohammad Zohaib. 2018. Dynamic Difficulty Adjustment (DDA) in Computer Games: A Review. *Advances in Human-Computer Interaction* (2018).