The Stress Is Real: Physiological Measurement of League of Legends Players Experience During a Live Esports Event

David Berga[1,2][a], Eleonora De Filippi[1][b], Arijit Nandi[1][c], Eulàlia Febrer Coll[3][d], Marta Reverté Bel[3], Lautaro J. Russo[4][e] and Alexandre Pereda-Baños[1][f]

[1]*Eurecat, Centre Tecnològic de Catalunya, Barcelona, Spain*

[2]*Elisava, Barcelona School of Design and Engineering (UVic-UCC), Spain*

[3]*eDojo - Electronic Dojo S.L., Barcelona, Spain*

[4]*Mental Gaming S.L., Barcelona, Spain*

dberga@elisava.net, {eleonora.defilippi, arijit.nandi, alexandre.pereda}@eurecat.org,

Keywords: Esports, League of Legends, Physiology, Arousal, Stress, Electrodermal Activity, Electrocardiogram, Photoplethysmogram, EDA, GSR, ECG, PPG, Heart Rate.

Abstract: Videogames and Esports experienced a huge growth in popularity lately and have opened a ripe new field for the study of human behavior. Esports gaming is an area in which videogame players need to cooperate and compete with each other, influencing their cognitive load, processing, stress, and as well as social skills. In this observational study we inquire whether variations in autonomic nervous system activity can be obtained reliably during a live League of Legends (LoL) event, especially considering this is a preliminary study with a limited participant sample. We found that game performance (winning or losing the game) significantly affects electrodermal activity and cardiac modulation, where players who lost the game showed higher stress-related physiological responses, compared to players who won. We also found that specific important events in the game, such as "Killing," "Dying," or "Destroying the turret," increased players' electrodermal and cardiac modulation compared with other less relevant events, such as "Placing the guards" or "Destroying the turret plates." Finally, by analyzing activity according to players' roles, we found various notable activity trends. Altogether, these (yet preliminary) results encourage further exploration of physiology-based applications for LoL and Esports players on live events.

1 INTRODUCTION

In Esports gaming video game players must cooperate and compete, often in highly demanding environments, affecting their cognitive load, emotional processing, stress levels, and social skills, among other aspects. Esports and videogames are on the rise, aided by the increasing ease with which online gaming environments allow interacting with other players, and they thus open a ripe area for neuroscience and experimental psychology research in more realistic and motivating settings than usually found in laboratory experiments (Pedraza-Ramirez et al., 2020). For example, technological developments involving wearable physiological, multimodal sensors, and interaction-based metrics, allow for close monitoring of cognitive emotional processes with high temporal resolution. This opens a landscape of possibilities not only for basic research, but also for research on advanced interaction techniques, such as the establishment of BCI (brain-computer-interface) loops that can be used for a wide variety of training, self-monitoring applications, and even to provide the audience with information on the state of the players. Of course, there is no question that games have become highly complex in the environments they represent, the skills they require, or the rules and roles involved, thus, any such applications would need to consider the specifics of the game under consideration.To this end we measured several metrics of electrodermal activity and cardiac modulation among competitive players dur-

[a] https://orcid.org/0000-0001-7543-2770
[b] https://orcid.org/0000-0001-5496-652X
[c] https://orcid.org/0000-0003-4238-5183
[d] https://orcid.org/0000-0003-1450-8086
[e] https://orcid.org/0009-0005-4224-6724
[f] https://orcid.org/0000-0002-1145-146X

Berga, D., De Filippi, E., Nandi, A., Coll, E., Bel, M., Russo, L. and Pereda-Baños, A.
The Stress Is Real: Physiological Measurement of League of Legends Players Experience During a Live Esports Event.
DOI: 10.5220/0012981700003828

In *Proceedings of the 12th International Conference on Sport Sciences Research and Technology Support (icSPORTS 2024)*, pages 199-205
ISBN: 978-989-758-719-1; ISSN: 2184-3201

ing several sessions of League of Legends games that took place in a live Esports event celebrated at a dedicated gaming venue. Together with the physiological metrics we obtained the activity logs from the corresponding games, provided by the game developer (Riot Games), to study if and how different events and roles elicit varying cognitive/emotional responses. aiming to ascertain the utility of such metrics for future biofeedback-based training of different game skills in LoL. Research in the field of psychology has traditionally focused on three main forms of both emotional and attentional responses: subjective perceptions of the individual about their own state, effects on behavior, and changes in their physiological patterns, such as acceleration or deceleration of heart rate and increase in skin conductivity (Bradley and Lang, 2000)(Mauss and Robinson, 2009). Each one of these approaches comes with both advantages and disadvantages. First, auto-informed methods, such as questionnaires or interviews, for which participants are directly asked to report their status are the only way to access the individual's subjective perception but are also limited by the individual's own ability to introspect, since many psychological processes can occur unconsciously or with low levels of consciousness (Nisbett and Wilson, 1977). Moreover, another limitation of this approach is that cognitive biases (such as social desirability bias) may interfere with the reports, thus making the information not entirely reliable. For this reason, in the field of experimental psychology research, the analysis of physiological responses (for example, variations in heart rate, skin conductivity or activation of facial muscles) has been introduced as a way of obtaining information about the cognitive and emotional processes of individuals in an indirect way. On the other hand, the main disadvantage of the physiological methods with respect to the self-reports concerns their ecological validity, so it is recommended to combine both methodologies (Cacioppo et al., 2017). Here, we used two main physiological measures:

Electrodermal Activity (or EDA, also known as Galvanic Skin Response or GSR) is a correlate of the activation of the sympathetic branch of the autonomic nervous system, which provides reliable information with a high temporal resolution about participants' physiological arousal, related to the intensity of experienced cognitive-emotional responses.

***Cardiac Modulations (ECG or Electrocardiogram),* or Photoplethysmography (PPG)** is a measure of cardiovascular pulse (i.e. Heart Rate). This signal is mediated by both the sympathetic and parasympathetic systems, thus responding to both physiological arousal and emotional regulation processes. This is dependent on both the intensity of emotions and their hedonic load, also with the appearance of cognitive resources for stimulus processing. The variability of heart rate can be analyzed by looking at the ratio of sympathetic vs parasympathetic modulation within the PPG signal.

Over the last decade, several physiological studies have been carried out in the Esports field (Kivikangas et al., 2011)(Argasiński and Grabska-Gradzińska, 2017)(Alhargan et al., 2017) and have reported relevant differences in participants' affective states during games. Among them, some studies have focused on the analysis of muscle signals (Electromyography/EMG) (Ahsan et al., 2009), of brain signals (Electroencephalography/EEG) (Hafeez et al., 2021), or of facial gestures (Samara et al., 2017). The main focus of these latter studies was to assess the emotional state of participants from these measures, classifying each of the 7 affective categories of basic emotions initially defined by (Ekman, 2005): anger, sadness, fear, disgust, joy, surprise, contempt, and neutral. Most of these studies have used psychophysiological tests such as the Russell test (Russell, 1980) or the Self Assessment Mannekin (Bradley and Lang, 1994), as a guideline for the validation of physiological measures that could affect the affective responses of participants. In a recent review by (Leis and Lautenbach, 2020), 17 studies were meta-analyzed in Esports contexts for psychological and physiological stress, and it was concluded that simply playing in an Esports non-competitive environment produced no stress reactions, whereas in competitive environments several studies reported increases in anxiety levels, cortisol levels, and physiological sympathetic activation, all three indicators of stress (Jones et al., 2012)(Yaribeygi et al., 2017). However, stress is not the only interesting indicator to consider in Esports environments, since peripheral physiology can also provide insight into various aspects of information processing, such as emotion, engagement, boredom or frustration.

In our study, we collected both EDA and ECG data while users were engaged in a widely played desktop videogame ”League of Legends”. The aim of this work was to evaluate the affective responses by analyzing distinct EDA and ECG metrics depending on game performance (winning or losing). Moreover, we investigated whether different game events (e.g. ”Killing”, ”Dying”, ”Kill Assist”, ”Destroy Turret”, ”Destroy Turret Plate” or ”Placing Ward”) elicited distinct physiological responses. Lastly, we wanted to elucidate whether different player roles (Jungle, Middle, Utility, Bottom, and Top) also impact physiological measures.

2 METHODS

2.1 Experimental Design

Sensors and Software. For our experiments we used a Shimmer3[1] pack with 5 simultaneous GSR+ Units providing Galvanic Skin Response for acquisition of Electrodermal Activity (EDA), as well as Cardiac Pulse (PPG) estimating heart rate variations. We developed our own Python tools (available in github[2]) for capturing data and sending it through Bluetooth (using pyserial and pylsl) and synchronizing that data with events of the game (using riotwatcher) for later statistical analysis for EDA (Ledalab[3] V3.4.9) and PPG (HeartPy[4]) respectively.

League of Legends Skills and Mechanics. League of Legends[5] is a Multiplayer Online Battle Arena (MOBA) videogame developed by Riot Games[6] in 2009. It has become one (if not) the most popular game of the current generation of online videogames [7] and has a significant impact on the economy of many other industrial and technology sectors (i.e. branding, consumables, social networks and media content). The original League of Legends game type mechanics (Summoner Rift) consists of a competition between 2 teams of 5 players in which each team has to destroy the enemy base or ”Nexus”. The battle zone is composed by 3 lanes (top, mid and bot) as well as a jungle (in between these lanes). Each team lane has 4 turrets and an inhibitor, and two more turrets in the Nexus. However, before reaching the enemy base, each team needs to destroy the turrets and the inhibitor of one of the enemy lanes. The usual teamwork consists of two players (Carry/Bottom and Utility/Support) controlling the bot lane, the Top/Tank controlling the top lane, the Middle controlling the mid lane and the Jungle that moves around all the lanes. As in other MOBAs, players will need to cooperate and sometimes play aggressively to kill the opponent’s champions and overcome the enemy positions. Riot users (summoners) have a specific experience level[8] (given its in-game time) and rank[9] (given its actual performance in ranked games).

Game Sessions and Subjects. A total of 4 sessions have been performed in the Asobu Esports Experience venue[10] with 12 participants (contacted and selected by United Gamers Academy[11]) in the gameplay experimentation. The participants were recruited through questionnaires reporting age, gender, skill level, rank level and game preferences. Only players with higher rank levels above silver III were selected for the study. Subjects’ age was between 18 and 25 years old (4 women and 12 men). Average player level was 216 (with lowest 82 and biggest 402) and corresponded to silver-gold S12 competitive rank gamers. Due to the limitation in the quantity of sensors available, we captured data from 12 participants on 4 sessions of playing a specific team during Summoner’s Rift gameplay (avg time 30-45 min), later filtered on 7 with enough valid events for statistical comparison. We cut the recording of these participants from the start to the end of the game and we set specific window times for each event (i.e. 5 sec). This data is synchronized with events downloaded from riotwatcher api[12]. Some events available for capture are ”killing”, ”dying”, ”kill assist”, ”special killing”, ”item purchased”, ”level up”, ”ward placed”, ”building kill”, ”champion transform”, ”turret plate destroyed” and ”elite monster kill”. After gameplay we annotated riot’s metadata for each participant such as game session data (total kills/deaths or damage done/received), win or loss condition and player roles (top, mid, bot, utility and jungle).

2.2 Physiological Data Processing

GSR Preprocessing. We have processed raw GSR data with Ledalab to extract the following measures: nrSCRs (total skin conductance number ”#” of responses above threshold), Latency (delay/surpassed time ”s” to elicit EDA with respect to the event), Amplitude (mean activity ”mV” inside the event window), PhasicMax (max phasic value ”mV” from the gap with respect the response and the event window)

[1] https://shimmersensing.com/

[2] https://github.com/dberga/riotwatcher-shimmer-pynput

[3] http://www.ledalab.de/

[4] https://pypi.org/project/heartpy/

[5] https://www.leagueoflegends.com/

[6] https://www.riotgames.com/en

[7] https://www.bcg.com/publications/2023/how-Esports-will-become-future-of-entertainment

[8] https://leagueoflegends.fandom.com/wiki/Experience_(summoner)

[9] https://leagueoflegends.fandom.com/wiki/Rank_(League_of_Legends)

[10] https://asobuEsports.com/

[11] https://unitedgamers.pro/

[12] https://developer.riotgames.com/

and Tonic (max tonic activity ”mV” with respect window). See Ledalab's documentation[13] for more details. Previous literature in electrodermal physiology has shown EDA can be a reliable quantifier of sympathetic dynamics (Posada-Quintero et al., 2016), meaning higher EDA correlated with higher sympathetic (stress/alert) levels.

The Matlab-based toolbox ”Ledalab” (Benedek and Kaernbach, 2010) was used for the GSR signal preprocessing and analysis. First, we carried out a preliminary visual examination to look for periodic drift in the signal, which reflects artifacts, and we resampled the raw signal to 50Hz using Neurokit2[14]. The following preprocessing operations were then carried out using Ledalab toolbox: low-pass Butterworth filtering with a cutoff frequency of 5 Hz, and smoothing to eliminate any remaining artifacts. Finally, we performed an event-related analysis utilizing the Continuous Decomposition Analysis (CDA) to extract the features indicating Skin Conductance Responses (SCRs). By extracting the phasic (driver) information underlying EDA, this approach attempts to obtain the signal features of the underlying sudomotor nerve activity. Skin conductance data is deconvoluted by the overall response shape, considerably enhancing temporal accuracy. This method enables the extraction of continuous phasic and tonic activity based on traditional deconvolution within a predetermined time window, which for us corresponded to a window comprising the three seconds before an event marker to the five following seconds. The number of SCRs within the response window, response latency for the first SCR, mean SCR amplitudes, maximum phasic, and average tonic activity within the specified window were therefore collected for each event described in the previous section.

PPG Data Processing. We have processed raw PPG data with Heartpy to obtain the BPM (”#” beats per minute), IBI (time ”ms” of interbeat interval or R-R), SDNN (standard deviation of intervals ”ms” between adjacent beats of the IBI of normal sinus beats), SDSD (standard deviation of successive differences between adjacent R-R intervals ”ms”) and RMSSD (root mean square of successive differences between adjacent R-R intervals ”ms”). The latter metrics (SDSD and RMSSD) are related to the measurement of HRV (heart rate variability). Indeed, higher HRV (higher values of SDSD or RMSSD) can represent parasympathetic/vagal modulation (associated with a state of relaxation), while a lower HRV (lower values for SDSD or RMSSD) represents sympathetic/flight-or-fight modulation (being stressed or alert; (Valenza et al., 2018)). Here we have to point out that studies on HRV are commonly analyzed over large timeline streams of heart rate data (about 5 min or more; (Shaffer and Ginsberg, 2017)) documentation explains the aforementioned metrics. However, our measurements of HRV are considering 5 to 10-second time windows according to the League of Legends fast-paced events.

Processing and analysis of raw PPG data were conducted using the Python-based toolkit ”Heartpy”(Van Gent et al., 2019), specialized for the analysis of PPG signal as compared to ECG. At every heartbeat, blood perfuses via the capillaries and arteries, causing the skin to become discolored. The PPG detects this discoloration. The systolic peak, diastolic notch, and diastolic peak make up the signal. First, as we did with the GSR signal, we resampled the raw PPG signal to 50Hz using Neurokit2. Then, we run the processing algorithm that comes with the Heartpy toolkit and which allows for the peak detection to extract reliable time-domain measures, such as beats per minute (BPM), and Interbeat Intervals (IBI). Furthermore, for each event, we extracted measures that reflect Heart Rate Variability (HRV) such as the RMSSD (root mean square of successive differences) and the SDSD (standard deviation of successive differences).

3 RESULTS

We performed data curation for our statistical analysis using data from 7 participants (a total of 2 game sessions with recorded measures in which 3 participants played twice) with enough event samples for later analysis and processing. Some of the data not mentioned in participants results was discarded for the final evaluation due to the incorrect samples from the sensor data and/or given the lack of relevant events in the gameplay to sync with the sensor data (e.g. given too much time being dead without interacting with game objects nor players). Here we select the players' sensor data collected with enough samples to retrieve individual PPG and EDA patterns to make valid statistical evaluations.

3.1 Physiological Results: Skin Conductance

In Table 1 we show mean statistics of nrSCR, Latency, Amplitude, PhasicMax, and Tonic values of players that win the gameplay and lose the gameplay. Similarly, in Table 2 we show statistics for

[13] http://www.ledalab.de/documentation.htm
[14] https://neuropsychology.github.io/NeuroKit/

Table 1: Win and Loss mean GSR metrics by stacking all events in one statistic. Observations from 7 participants playing during 2 sessions. nrSCR: skin conductance responses over threshold. *$p < 0.05$ between all events.

Result	nrSCR	Latency	Amplitude	PhasicMax	Tonic
WIN	1.85±1.73	-0.15±1.93	0.27±0.63	0.55±1.17	12.58±7.05
LOSS	2.60±2.21	-0.76±1.70	0.19±0.41	0.37±0.55	6.88±4.74
TOTAL	2.37±2.10	-0.57±1.79	0.22±0.49	0.42±0.79	8.61±6.12

Table 2: Event mean GSR metrics from events "Killing", "Dying", "Placing Ward", "Destroying Turret" and "Destroying Turret Plate". Observations from 7 participants playing during 2 sessions. nrSCR: skin conductance responses over threshold. *$p < 0.05$ between win/lose.

EVENT	nrSCR	Latency	Amplitude	PhasicMax	Tonic	N
KILL	1.23±1.17	0.05±2.11	0.03±0.04	0.10±.07	7.76±3.23	9
DIE	2.42±2.10	-0.19±0.74	0.32±0.63	0.50±.86	12.20±5.82	35
PLACE WARD	2.32±2.16	-0.30±2.15	0.23±0.45	0.15±0.14	12.06±5.41	75
DES.TURRET	2.18±2.21	-0.61±1.54	0.16±0.23	0.25±0.44	9.75±6.27	79
DES.PLATE	2.18±1.89	-0.46±1.84	0.29±0.59	0.46±0.68	3.09±1.77	49

events "Killing", "Dying", "Place Ward", "Destroy Turret" and "Destroy Turret Plate". We expand these statistics in Supplementary Material-Table 6 filtering player roles in the game.

Given the Chi-squared measured distributions (non-parametric) we performed Friedman's tests over win-loss and event data for each GSR metric. On analyzing winning or losing the match (Tables 1-2), we observed that the nrSCR, Amplitude, and PhasicMax activities were significantly higher during the "Killing" event for players (p = .046, χ^2 = 4.000). Additionally, nrSCR and Amplitude values were also significantly elevated during the "Destroying Turret" event (p = .020, χ^2 = 5.444), while nrSCR activity alone showed a significant increase during the "Destroying Plate" event (p = .008, χ^2 = 7.143), with Amplitude showing a trend towards significance (p = .071, χ^2 = 3.266). Tonic activity differed significantly only in relation to the "Dying" event (p = .035, χ^2 = 4.455) and "Placing Ward" event (p = .002, χ^2 = 10.000) between winning and losing conditions.

When comparing GSR activity distributions across all events for winning players, we found that Latency was significantly shorter (p = .024, χ^2 = 11.265), Amplitude was significantly higher (p = .041, χ^2 = 9.959), and Tonic activity was significantly greater (p = .010, χ^2 = 13.28). PhasicMax showed a trend towards higher values, though it did not reach statistical significance (p = .092, χ^2 = 8.000). In contrast, there were no significant differences in GSR activity across events for players who lost the game.

3.2 Physiological Results: Heart Rate

In Table 3 and Supplementary Material-Table 5 we show mean statistics of pulse metrics according to win condition, event, and role.

After performing Friedman tests over PPG metrics for all events, we found that SDSD was significantly higher when winning the game (p = .016, χ^2 = 10.371). BPM was significantly lower and IBI was significantly longer (p = .041, χ^2 = 8.28). We also tested for differences between winning and losing the game for each specific event. For "Destroying Turret Plate", SDSD was significantly higher (p = 5.32×10^{-4}, χ^2 = 12.0) and RMSSD was significantly increased (p = .004, χ^2 = 8.333). During the "Dying" event, SDSD was significantly elevated (p = .011, χ^2 = 6.4) and RMSSD was also higher (p = .002, χ^2 = 10). Additionally, SDSD showed a significant increase when "Placing a Ward" (p = .011, χ^2 = 6.4).

4 CONCLUSIONS

This study shows the potential of using physiological measurements (EDA and ECG) to monitor cognitive and emotional processes in complex game environments such as League of Legends, and supports the idea that these metrics could enable biofeedback based loops for interaction, training, and showcasing purposes. In the study, we characterized physiological responses depending on performance, events as well as participants' roles in the game.

Table 3: Win and Loss mean PPG metrics by stacking all events in one statistic. Observations from 7 participants. BPM: Beats per minute; IBI: interbeat interval; SDNN: deviation between adjacent beats; SDSD: deviation of differences between R-R intervals; RMSSD: successive differences between R-R intervals. $*p < 0.05$ between all events.

Result	BPM	IBI	SDNN	SDSD	RMSSD
WIN	172±113	531±333	129±48	*100±47	202±87
LOSS	*94±50	*743±212	77±52	57±44	117±92
TOTAL	114±80	688±265	90±56	68±49	140±98

Table 4: Event mean PPG metrics from events ”Killing”, ”Dying”, ”Placing Ward”, ”Destroying Turret” and ”Destroying Turret Plate”. Observations from 7 participants playing during 2 sessions. BPM: Beats per minute; IBI: interbeat interval; SDNN: deviation between adjacent beats; SDSD: deviation of differences between R-R intervals; RMSSD: successive differences between R-R intervals. N are event occurrences. $*p < 0.05$ between win/lose.

EVENT	BPM	IBI	SDNN	SDSD	RMSSD	N
KILL	69±10	892±139	73±46	57±31	99±54	7
DIE	103±66	726±249	87±54	*67±46	*132±86	37
PLACE WARD	122±75	648±279	92±54	*76±54	143±94	31
DES.TURRET	118±95	678±260	95±56	72±52	146±96	57
DES.PLATE	117±78	673±276	*90±60	*64±47	140±111	71

In most cases, we found significant differences in EDA (for nrSCR, Amplitude, and PhasicMax activity) during ”Killing”, ”Destroying Turret” or ”Destroying Turret Plate” between players that are winning the game and players that are losing the game, by showing more relaxed states for winning players. One should consider that more relaxed state cannot be achieved when there is greater activity on GSR. Moreover, when players were winning the game, they showed distinct patterns of physiological activity depending on the events in the game (e.g., ”Killing”, ”Destroying Turret”, ”Destroying Plate”). In contrast, we did not find any significant difference between these events for players that were losing the game, as they remained overall with lower HRV modulation. This can hinder the possibility that players that perform badly show similar physiological states (being alert and/or excited) across the game, while players that perform well have distinct physiological behavior during the course of game events.

For the case of PPG, similarly to the aforementioned, SDSD was significantly distinct for players that were winning the game between different events. On the other side, we found that only IBI and BPM measures showed significant differences for players that were losing the game. Overall, players that performed better (winning) showed significantly higher parasympathetic modulation (i.e., relaxation) than the ones that were losing. These results suggest that poor game performance induces higher stress or to be in a state of alert, while players that perform better tend to remain in more relaxed states. Furthermore, the analysis for specific events, like ”Dying”, ”Destroying Turret Plate” or ”Placing Ward”, has shown that players have distinct values of SDSD and RMSSD, with Killing” or ”Dying” events inducing higher sympathetic modulation (lower HRV).

Study Limitations and Future Work. Despite the lack of physiological samples for participants and game sessions we obtained enough measurements to pinpoint differences in-game performance and events, confirming the potential for using these measures in live events. By having a higher number of participants and game sessions we would suggest undergoing similar studies, not only for analyzing physiology over game performance and events but also for conducting an in-depth analysis of game roles, champions, player level, and type of match (beyond League of Legends' summoner's rift) in relation with EDA and ECG measurements. Further exploration in this context would expand both research and development capabilities on evaluating physiological responses in eSports.

4.1 Funding

This study has been possible through the Grant IRC 2020 (Reference ACE033/21/000046) funded by ACCIÓ (Catalan Agency for Business Competitiveness), from the project ”Esports-LAB” lead by INDESCAT (Associació Catalana Clúster de la Indústria de l'Esport), partners with Generalitat de Catalunya, EsportCat and Fundació Catalana per l'Esport.

4.2 Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. Experiment participants signed an authorization form prior to the study to authorize the usage of the captured physiological data as well as performance from their Riot Gamertag and remained anonymous according to the Spanish national law LOPD (Ley Orgánica de Protección de Datos de Carácter Personal).

REFERENCES

Ahsan, M. R., Ibrahimy, M. I., and Khalifa, O. O. (2009). Emg signal classification for human computer interaction: a review. volume 33, pages 480–501.

Alhargan, A., Cooke, N., and Binjammaz, T. (2017). Multimodal affect recognition in an interactive gaming environment using eye tracking and speech signals. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*. ACM.

Argasiński, J. K. and Grabska-Gradzińska, I. (2017). Patterns in serious game design and evaluation application of eye-tracker and biosensors. In *Artificial Intelligence and Soft Computing*, pages 367–377. Springer International Publishing.

Benedek, M. and Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of neuroscience methods*, 190(1):80–91.

Bradley, M. M. and Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59.

Bradley, M. M. and Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology. In Lane, R. D. R., Nadel, L., Ahern, G. L., Allen, J., and Kaszniak, A. W., editors, *Cognitive Neuroscience of Emotion*, pages 25–49. Oxford University Press.

Cacioppo, J. T., Chen, H. Y., and Cacioppo, S. (2017). Reciprocal influences between loneliness and self-centeredness: A cross-lagged panel analysis in a population-based sample of african american, hispanic, and caucasian adults. *Personality and Social Psychology Bulletin*, 43(8):1125–1135.

Ekman, P. (2005). Basic emotions. In *Handbook of Cognition and Emotion*, pages 45–60. John Wiley & Sons, Ltd.

Hafeez, T., Saeed, S. M. U., Arsalan, A., Anwar, S. M., Ashraf, M. U., and Alsubhi, K. (2021). EEG in game user analysis: A framework for expertise classification during gameplay. *PLOS ONE*, 16(6):e0246913.

Jones, B. J., Tan, T., and Bloom, S. R. (2012). Minireview: Glucagon in stress and energy homeostasis. *Endocrinology*, 153(3):1049–1054.

Kivikangas, J. M., Chanel, G., Cowley, B., Ekman, I., Salminen, M., Järvelä, S., and Ravaja, N. (2011). A review of the use of psychophysiological methods in game research. *Journal of Gaming & Virtual Worlds*, 3(3):181–199.

Leis, O. and Lautenbach, F. (2020). Psychological and physiological stress in non-competitive and competitive esports settings: A systematic review. *Psychology of Sport and Exercise*, 51:101738.

Mauss, I. B. and Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & Emotion*, 23(2):209–237.

Nisbett, R. E. and Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, 35(4):250–256.

Pedraza-Ramirez, I., Musculus, L., Raab, M., and Laborde, S. (2020). Setting the scientific stage for esports psychology: a systematic review. *International Review of Sport and Exercise Psychology*, 13(1).

Posada-Quintero, H. F., Florian, J. P., Orjuela-Cañón, A. D., and Chon, K. H. (2016). Highly sensitive index of sympathetic activity based on time-frequency spectral analysis of electrodermal activity. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 311(3):R582–R591.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178.

Samara, A., Galway, L., Bond, R., and Wang, H. (2017). Affective state detection via facial expression analysis within a human computer interaction context. *Journal of Ambient Intelligence and Humanized Computing*, 10(6):2175–2184.

Shaffer, F. and Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in Public Health*, 5.

Valenza, G., Citi, L., Saul, J. P., and Barbieri, R. (2018). Measures of sympathetic and parasympathetic autonomic outflow from heartbeat dynamics. *Journal of Applied Physiology*, 125(1):19–39.

Van Gent, P., Farah, H., Van Nes, N., and Van Arem, B. (2019). Heartpy: A novel heart rate algorithm for the analysis of noisy signals. *Transportation research part F: traffic psychology and behaviour*, 66:368–378.

Yaribeygi, H., Panahi, Y., Sahraei, H., Johnston, T. P., and Sahebkar, A. (2017). The impact of stress on body function: a review. *EXCLI Journal; 16:Doc1057; ISSN 1611-2156*.

SUPPLEMENTARY MATERIAL

See Tables 5 and 6 in https://drive.google.com/file/d/1zByOo59gS2x0cGhZLhY5akaVFhTGlAQi/view?usp=sharing.