2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)
Toronto, Canada, June 24-28, 2019

Using Physiological Linkage for Patient State Assessment In a Competitive Rehabilitation Game

Ali Darzi[1], Domen Novak[1]

[1]Department of Electrical and Computer Engineering, University of Wyoming, Laramie, WY 82071, USA,

{adarzi, dnovak1}@uwyo.edu

***Abstract*— Competitive rehabilitation games can enhance motivation and exercise intensity compared to solo exercise; however, since such games may be played by two people with different abilities, game difficulty must be dynamically adapted to suit both players. State-of-the-art adaptation algorithms are based on players' performance (e.g., score), which may not be representative of the patient's physical and psychological state. Instead, we propose a method that estimates players' states in a competitive game based on the covariation of players' physiological responses. The method was evaluated in 10 unimpaired pairs, who played a competitive game in 6 conditions while 5 physiological responses were measured: respiration, skin conductance, heart rate, and 2 facial electromyograms. Two physiological linkage methods were used to assess the similarity of the players' physiological measurements: coherence of raw measurements and correlation of heart and respiration rates. These linkage features were compared to traditional individual physiological features in classification of players' affects (enjoyment, valence, arousal, perceived difficulty) into 'low' and 'high' classes. Classifiers based on physiological linkage resulted in higher accuracies than those based on individual physiological features, and combining both feature types yielded the highest classification accuracies (75% to 91%). These classifiers will next be used to dynamically adapt game difficulty during rehabilitation.**

Keywords*— rehabilitation, serious games, affective computing, physiological linkage, coregulation, psychophysiology

I. Introduction

Serious games have become a popular rehabilitation tool, as they result in higher motivation, higher exercise intensity, and better rehabilitation outcome than traditional therapy [1]. While such rehabilitation games traditionally involved a single patient, newer 'interpersonal' rehabilitation games allow patients to compete or cooperate with another person: another patient, a therapist, or an unimpaired friend or relative [2], [3]. As competition and cooperation are important sources of intrinsic motivation, such interpersonal games have the potential to further increase motivation and exercise intensity compared to single-patient rehabilitation games, as demonstrated by preclinical studies [4]–[8]. However, since such games often involve two players with different skills and abilities (e.g., a patient and a therapist or a patient and a loved one), intelligent difficulty adaptation algorithms are needed to ensure that the game is appropriately difficult for both players.

Research supported by the National Science Foundation under grant no. 1717705 as well as by the National Institute of General Medical Sciences of the National Institutes of Health under grant no. P20GM103432.

In single-player rehabilitation games, difficulty adaptation is commonly done with two approaches: task performance and/or physiological measurements. Two examples of task performance are in-game score [9] and the number of successfully completed actions [10], which can be easily measured and used to adapt difficulty with if-then rules (e.g., if score is too low, make the game easier). Although task performance does provide an indicator of the patient's capabilities, it does not provide information about subjective states such as workload; for example, a patient can be successfully performing the exercise but only at the cost of extremely high workload that is likely to lead to frustration and fatigue. For this reason, some researchers have used physiological measurements such as heart rate (HR), respiration rate (RR) and electrodermal activity (EDA) to estimate the patient's state and adapt difficulty accordingly (e.g., if patient workload is high, reduce game difficulty) [11], [12]. However, as physiological measures require additional hardware, signal processing and classification methods, they are not used as commonly as simple performance measures despite the potential to provide additional information.

Competitive and cooperative rehabilitation games have almost exclusively focused on performance-based difficulty adaptation methods [4]–[8]. In our previous paper, we preliminarily evaluated a physiology-based adaptation method for a competitive rehabilitation game, but that method only adapted difficulty to suit one player based on that player's physiological responses [13]. Nonetheless, it demonstrated the feasibility of physiology-based difficulty adaptation for competitive and cooperative rehabilitation games. In the current paper, our goal is to measure the physiological responses of both players and use them as a basis for player state estimation in a competitive game – for example, to say whether one player is enjoying the game more than the other. This estimated state can then be used as a basis for game difficulty adaptation.

In single-player rehabilitation games, difficulty is adapted based on features extracted from the patient's physiological responses – e.g., heart rate, respiration rate [11], [12]. In a competitive rehabilitation game, we could similarly adapt difficulty based on, e.g., the heart rate and respiration rate of each individual player. However, there is an additional possible source of information: physiological linkage, a concept that indicates the degree of synchronization (i.e., mutual variation) of physiological responses between two people. Physiological linkage has been shown to occur during both competitive and cooperative tasks; furthermore, the degree of linkage increases with the amount of cooperation [14]–[16] or the intensity of competition [17], [18]. Thus, physiological linkage could potentially provide insights into competition and cooperation that cannot be gleaned by simply analyzing each participant's physiological responses in isolation, as suggested by authors in other areas of human-machine interaction [19].

This paper presents, to the best of our knowledge, the first evaluation of two participants' physiological responses in a competitive rehabilitation game. Each individual participant's physiological responses as well as different metrics of physiological linkage are evaluated in different game states: single-player (one participant plays and the other one watches), fair and slow competitive, fair and fast competitive, and unfair competitive. Our research questions were:

RQ1: Can a combination of individual physiological responses and physiological linkage be used to classify participants' affects (enjoyment, emotional valence and arousal, perceived difficulty) in a competitive rehabilitation game? This would allow us to use physiological responses to dynamically adapt the difficulty of competitive rehabilitation games.

RQ2: Does physiological linkage provide information that cannot be obtained by simply analyzing both participants' individual physiological responses in isolation? As physiological linkage is a relatively new metric that requires more advanced signal processing than individual physiological responses, it should only be used in practical systems if it provides additional useful information.

II. Methodology

A. Study Setup

The study involved 10 pairs of healthy university students (26.8 years old, 7 females and 1 left-handed) who experienced different conditions of a competitive arm rehabilitation game that was reused from a previous study by our research group [5]. It is a Pong game consisting of two paddles and a puck on a board (Fig. 1). Each participant moves a paddle left and right using a commercial joystick (3D Pro, Logitech) and tries to block the puck with their paddle. If the puck passes a player's paddle and reaches the top or bottom of the screen, the other player scores a point and the puck is instantly moved to the middle of the board, where it remains stationary for a second before moving in a random direction. The game difficulty can be adjusted using the ball speed and the size of the paddles.

Two g.USBamp measurement amplifiers and associated sensors (g.tec Medical Engineering GmbH, Austria) were placed behind the participants in order to measure five physiological responses for each participant (Fig. 1, left). The electrocardiogram (ECG) was recorded using four electrodes on the body (two on the chest, one on the back over the spine, and one on the abdomen) as recommended by the manufacturer of the g.USBamp. Respiration was recorded using a thermistor-based flow sensor (g.tec) in front of the nose and mouth. EDA was recorded via two electrodes (g.GSRsensor2, g.tec) attached to the fingertips of the index and middle fingers of the non-dominant hand (Fig.1, bottom right). Two facial electromyograms (FEMG) were recorded using 5 electrodes: two on the zygomaticus major muscle to measure smiling, two on the corrugator supercilii to measure frowning, and a ground electrode on the forehead (Fig. 1, top right). An experimenter sat on the participants' right side to carry out the study protocol and oversee the measurements.

B. Study Protocol

The study protocol started with a detailed explanation of the study purpose and procedure; after that, participants signed an informed consent form. The physiological sensors were attached, and physiological responses were measured for a 3-minute baseline period, during which participants did not do anything and were instructed to relax. The main part of experiment then consisted of six conditions, each three minutes long (Fig. 2). The six conditions consisted of two single-player conditions (one for each player), one fair and slow competitive condition, one fair and fast competitive conditions, and two unfair conditions. During the single-player conditions, the speed and paddle size were set to a medium setting (identical to that used in our previous study [5] to provide a challenging but not frustrating experience for the average participant), and one participant played against a computer-controlled opponent while the other participant watched. In both fair competitive conditions, the paddle size was medium for both players. In the 'fair and slow' condition, the speed was 70% of the medium value; in the 'fair and fast' condition, the speed was 130% of the medium value. In the unfair competitive conditions, the speed was set to medium, but one participant played with a small paddle while the other played with a large paddle (50% or 150% of the width of the medium paddle). To experience the unfair condition, each participant played with the small paddle once.

Five physiological measurements (ECG, respiration, EDA, 2 FEMG) were recorded during each condition, and participants were instructed to not speak in order to reduce noise in FEMG and respiration signals. However, they were allowed to make noises, laugh and frown during the conditions. After each condition, a short questionnaire was filled out to assess perceived difficulty (1-7), enjoyment (1-

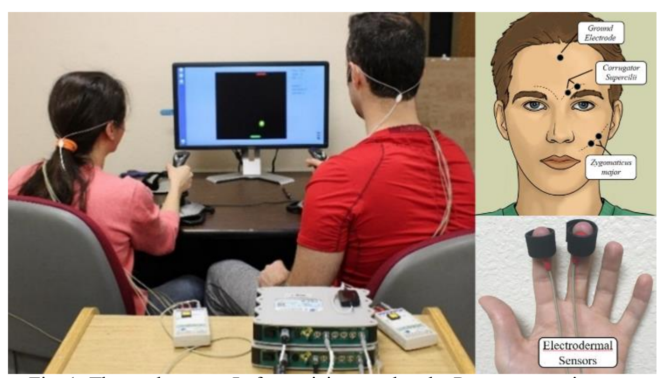

Fig. 1. The study setup. Left: participants play the Pong game using two joysticks while physiological responses are recorded with two identical recording devices. Top right: placement of the facial electromyogram electrodes (adapted from wikihow.com under Creative Commons License). Bottom right: a close-up of the skin conductance electrodes.

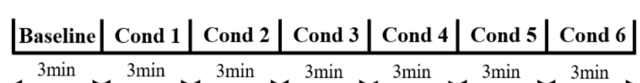

Fig. 2. The study protocol: a baseline period followed by 6 conditions, each 3 minutes long. Cond: condition.

7), valence (1-9, with 1 being very negative and 9 being very positive), and arousal (1-9). The conditions were played in random order.

C. Individual Physiological Features

Five physiological measurements were recorded from each participant during the baseline and six game conditions with a sampling frequency of 256 Hz. For each condition, 19 physiological features were calculated from each participant's five measurements.

ECG: Two time-domain features were calculated: mean HR, and standard deviation of inter-beat intervals. Furthermore, three frequency-domain features of HR variability were calculated: power of low frequencies (LF), power of high frequencies (HF) and the power ratio of LF/HF. The LF range is 0.04–0.15 Hz while the HF range is 0.15–0.4 Hz [20].

EDA: Five features were calculated from tonic (low-frequency) and phasic (high-frequency) components of EDA: the mean EDA, EDA difference between the first and last second of the condition, number of skin conductance responses in the condition, the mean response amplitude, and the standard deviation of response amplitude [21].

Respiration: The mean RR (number of complete breathing cycles per minute), the standard deviation of RR, and the root-mean-square of successive differences of respiration periods were calculated.

FEMG: Three features were calculated from each FEMG measurement. First, the root mean square (RMS) value was calculated over the entire condition. Then, the RMS value of the signal was calculated over consecutive 5-second windows within the condition, and the maximum and minimum RMS value over all windows were calculated as the second and third features. As smiling and frowning can be very brief, they might be missed by the RMS value over the entire condition.

All calculations were done both with normalization and without normalization. To normalize the data, the values for the six conditions were divided by the baseline value.

D. Physiological linkage

Two physiological linkage methods were selected from the literature [22], [23]: the coherence of raw measurements and the correlation of instantaneous HR and RR signals.

Coherence of raw physiological measurements: Welch's averaged periodogram method was used to calculate the coherence between two equivalent raw measurements (e.g. EDA of participant 1 and participant 2). Coherence is a standard signal processing method that finds the co-oscillation of two measurements up to half of the sampling frequency [24]. While some of the signals are not expected to contain useful data up to 128 Hz, we nonetheless calculated coherence up to 128 Hz for all five equivalent signals in order to be consistent.

Correlation of instantaneous HR and RR: While raw ECG and respiration signals are not expected to correlate between participants, the extracted HR and RR might be better options. To study this possibility, instantaneous HR and RR were calculated from peak-to-peak intervals in each condition. Then, linear interpolation was used to create a 256-Hz signal from the obtained instantaneous HR and RR values. Finally, physiological linkage was calculated as the Pearson correlation coefficient between the two participants' interpolated instantaneous HR and RR. This resulted in two Pearson coefficients for each condition: one for HR and one for RR.

Calculations of coherence were done both with and without normalization. Correlations of instantaneous HR and RR were not normalized since they already have a range of -1 to 1.

E. Statistical Analysis

One-way analyses of variances (ANOVA) followed by Tukey post-hoc tests were used to identify the physiological individual and linkage features that differed significantly between study conditions. The significance threshold was set to $p = 0.1$ due to the exploratory nature of the study.

F. Classification and validation

In a rehabilitation game, the ultimate goal is to classify the participant's affects and use them as a basis for game difficulty adaptation. While we did not perform closed-loop adaptation in this study, we did use physiological measurements to classify four affects into two possible classes (low/high) for each recorded condition. The four affects were enjoyment, emotional valence, arousal, and perceived difficulty; all were obtained from the short questionnaire (section II.B). As classification combines multiple features, it may be accurate even if no single feature shows significant differences between conditions (section II.E).

As we were interested in comparing the classification performance of the individual physiological features and the physiological linkage features, classification was performed using three different input datasets: individual features only, linkage features only, and both types of features together. The input data was classified into either 'low' or 'high' classes based on the short questionnaire: 'low' was an answer of 1-3 for all affects while 'high' was an answer of 5-7 for enjoyment or perceived difficulty and 7-9 for emotional valence or arousal. Classification of each of the four affects was done independently. Conditions where participants' responses did not fall into either 'low' or 'high' classes (e.g., answer "4") were omitted, as we wished to limit ourselves to situations with clearly expressed affects.

Classification of each affect began with stepwise regression-based feature selection [25] (threshold: $p = 0.1$) that selected the most relevant set of input features. Then, two different classifiers (support vector machine (SVM) with a linear kernel, and ensemble decision tree (EDT)) were used to classify the selected features into 'low' and 'high' classes for each recorded condition. The classifiers were validated using 10-fold crossvalidation (90% of conditions used to train, 10% to test the classifier; procedure repeated 10 times with different test data).

III. Results

A. Short Questionnaire Results

Table I shows the mean values of four affects for the first and second participants across all ten pairs. It demonstrates that different affects were successfully induced in the different conditions. For example, the highest perceived difficulty and highest arousal were observed in the unbalanced conditions. The ‘fair and fast’ game condition induced a high difficulty level for both participants, but also induced the highest enjoyment level and the most positive emotional valence.

B. *Individual Physiological Features*

Table II shows the individual physiological features that varied significantly between conditions. For each feature, the table also lists the ‘most different condition’ as the condition that had the largest number of significant differences from other conditions in post-hoc tests. As the conditions result in different affects (Table I), the features in Table II can be considered as the most informative for discrimination of different affective states. As an example, Fig. 3 shows the RR of both participants in different conditions. The participants’ RR was lower in the baseline period than the other six conditions, which is also indicated in Table II.

C. *Physiological Linkage Features*

1) *Coherence of Raw Physiological Measurements*

Table III shows the physiological linkage features obtained from coherence of raw physiological measurements that varied significantly between study conditions. For each feature, the table also lists the ‘most different condition’ (defined as in III.B). Nonnormalized EDA and FEMG of frowning mainly showed differences between baseline and the other conditions, which is not necessarily informative; however, normalized FEMG of frowning showed that the 2-player fast and fair condition differs from the others. This normalized coherence feature of FEMG of frowning is shown in Fig. 4.

2) *Correlation of Heart and Respiration Rates*

Fig. 5 presents the correlation of instantaneous HR and RR for the baseline and 6 game conditions of 10 pairs. The instantaneous HR correlation shows an increase for ‘2-player fair and fast’ condition compared to other conditions, especially the single-player ones. The instantaneous RR correlation coefficient is highest in the ‘2-player fair and fast’ and ‘2-player unfair to participant 1’ conditions.

TABLE I. MEAN OF FOUR AFFECTS OBTAINED FROM THE SHORT QUESTIONNAIRE ACROSS ALL TEN PARTICIPANT PAIRS (SP-S1: 1-PLAYER OF PARTICIPANT 1, SP-S2: 1-PLAYER OF PARTICIPANT 2, 2P-F-S: 2-PLAYER FAIR AND SLOW, 2P-F-F: 2-PLAYER FAIR AND FAST, 2P-U1: 2-PLAYER UNFAIR TO PARTICIPANT 1, 2P-U2: 2-PLAYER UNFAIR TO PARTICIPANT 2).

	Participant	SP-S1	SP-S2	2P-F-S	2P-F-F	2P-U1	2P-U2
Difficulty	1	3.6	2.2	3.2	4.8	6.1	2.1
	2	2.1	3.5	2.5	5.0	2.4	5.9
Enjoyment	1	4.2	2.8	3.7	5.3	4.8	4.5
	2	3.8	4.3	4.4	5.2	4.7	5.4
Valence	1	6.0	4.5	5.4	6.8	5.9	5.9
	2	6.3	6.6	6.9	7.3	6.7	7.2
Arousal	1	4.1	2.8	4.1	5.5	6.3	4.0
	2	4.1	4.7	4.8	6.4	3.7	6.8

TABLE II. THE MOST INFORMATIVE INDIVIDUAL FEATURES SELECTED BY ANALYSIS OF VARIANCE. THE MOST DIFFERENT CONDITION WAS THE ONE WITH LARGEST NUMBER OF SIGNIFICANT DIFFERENCES IN THE POST-HOC TESTS. RMS: ROOT MEAN SQUARE, EDA: ELECTRODERMAL ACTIVITY, FEMG: FACIAL ELECTROMYOGRAM.

	Participant	P-value	Most Different Condition
Respiration Rate	1	0.004	Baseline
	2	0.001	Baseline
Heart Rate	1	0.004	2P-F-F
	2	0.001	SP-S2
RMS of Smiling FEMG	1	0.001	2P-F-F
	2	0.001	2P-F-S
Mean EDA amplitude	1	0.001	2P-U2
	2	0.001	SP-S2

TABLE III. THE MOST INFORMATIVE PHYSIOLOGICAL LINKAGE FEATURES (COHERENCE OF RAW PHYSIOLOGICAL MEASUREMENTS). P-VALUE THRESHOLD = 0.1, 2P-F-F: 2-PLAYER FAIR AND FAST, EDA: ELECTRODERMAL ACTIVITY, FEMG: FACIAL ELECTROMYOGRAM

	Normalized	Measurement	P-Value	Most different condition
1	×	EDA	0.0775	Baseline
2	×	FEMG Frown	0.0035	Baseline
3	✓	FEMG Frown	0.0980	2P-F-F

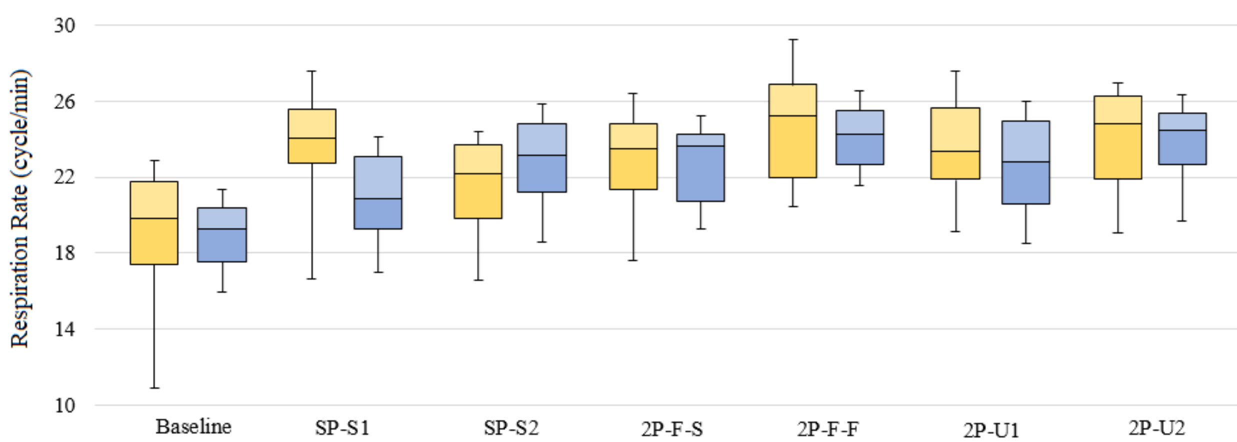

Fig. 3. The nonnormalized respiration rate for seven conditions. The data of first and second participants are shown in orange and blue colors, respectively. SP-S1: 1-player of participant 1, SP-S2: 1-player of participant 2, 2P-F-S: 2-player fair and slow, 2P-F-F: 2-palyer fair and fast, 2P-U1: 2-player unfair to participant 1, 2P-U2: 2-player unfair to participant 2.

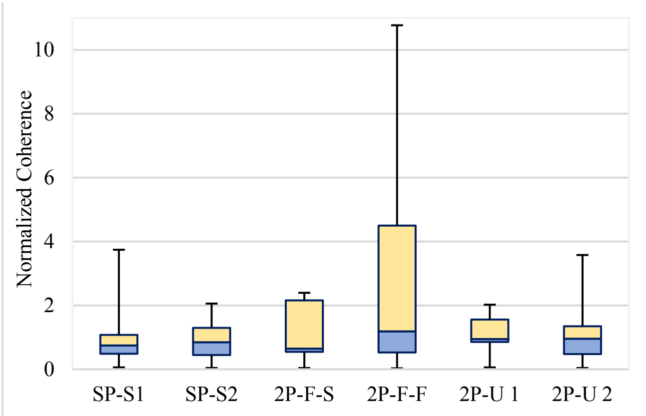

Fig. 4. Normalized coherence (a dimensionless variable) for the electromyogram of the corrugator supercilii across 10 pairs. SP-S1: 1-player of participant 1, SP-S2: 1-player of participant 2, 2P-F-S: 2-player fair and slow, 2P-F-F: 2-player fair and fast, 2P-U1: 2-player unfair to participant 1, 2P-U2: 2-player unfair to participant 2.

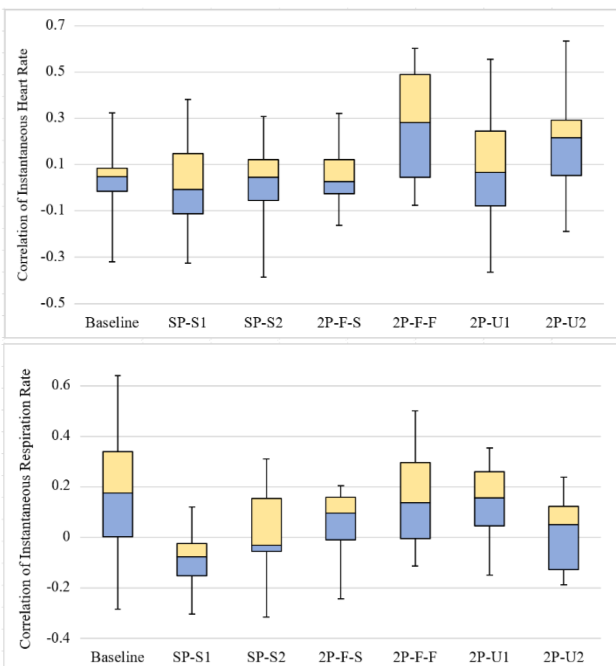

Fig. 5. Box and whisker plots for instantaneous heart rate correlation (top) and respiration rate correlation (bottom). Calculated as Pearson coefficient (range: -1 to 1). SP-S1: 1-player of participant 1, SP-S2: 1-player of participant 2, 2P-F-S: 2-player fair and slow, 2P-F-F: 2-player fair and fast, 2P-U1: 2-player unfair to participant 1, 2P-U2: 2-player unfair to participant 2.

D. *Classification of Four Affects*

Table IV presents the classification accuracy of both participants' four affects using different feature types as the inputs. In most classification cases, physiological linkage features yielded higher accuracies than individual physiological features. However, the highest classification accuracy was almost always obtained using the combination of both feature types. Among the two classification methods, SVM showed higher overall accuracy. The highest classification accuracy (over 90% for two-class classification) was obtained for the second participant's emotional valence while the lowest accuracy was obtained for the first participant's perceived difficulty (75.1%).

TABLE IV. CLASSIFICATION ACCURACY OF FOUR AFFECTS OF FIRST AND SECOND PARTICIPANTS. THE CLASSIFIER USED WAS EITHER A SUPPORT VECTOR MACHINE (NO MARKING) OR AN EMSEMBLE DECISION TREE ($^{\text{EDT}}$).

Affect	Participant	Individual Physiological Features	Linkage Features	Both
Perceived difficulty	1	65.3%	66.3%	75.1%
	2	74.3%	68.4%	76.3%
Enjoyment	1	70.0%	68.8%	82.5%
	2	80.2%	$^{\text{EDT}}$ 85.2%	83.3%
Valence	1	$^{\text{EDT}}$ 65.0%	72.5%	76.5%
	2	$^{\text{EDT}}$ 81.2%	76.5%	90.8%
Arousal	1	79.0%	83.6%	80.4%
	2	$^{\text{EDT}}$ 66.8%	74.8%	78.3%

IV. Discussion

A. *Feasibility of Affect Classification*

Our results show that physiological responses can be used to classify the affective state of two participants in a competitive arm rehabilitation game, with two-class accuracies ranging from 75% to 91%. While the target accuracy for a rehabilitation game is unclear, studies of physiology-based affect classification in entertainment games suggest that 80-90% is sufficient for an enjoyable user experience [26]. Furthermore, physiological linkage increased accuracy by up to 10% compared to using only individual features (Table IV). Studies of entertainment games suggest that increasing accuracy from 70% to 80% results in a measurable increase in user enjoyment [26], and we thus believe that physiological linkage is worth analyzing despite the additional necessary signal processing.

As the next step, we will connect our classifiers to decision-making rules that will adapt the difficulty of the competitive game. For example, if one player is experiencing high perceived difficulty while the other is not, difficulty will be increased for the "better" player only (by, e.g., making their paddle smaller). Since the classification is based on 3 minutes of data, difficulty can be adapted every 3 minutes, which is similar to adaptation intervals used in single-player rehabilitation games [11-12]. Such physiology-based adaptation may allow more effective difficulty adaptation in competitive rehabilitation games than is achieved with performance-based methods, resulting in higher motivation and exercise intensity. Furthermore, the same approach could also be generalized to competitive and cooperative scenarios outside rehabilitation, such as collaborative learning.

Admittedly, our study only involved unimpaired participants. As stroke survivors have different physiological responses to rehabilitation exercises than unimpaired participants [27], they also might not exhibit physiological linkage with other stroke survivors. Thus, future studies in this area should recruit stroke survivors and other patient groups to validate the clinical feasibility of the proposed approach. Furthermore, our study did not evaluate the clinical practicality of the approach: patients may not enjoy wearing the physiological sensors, and the sensor setup time may reduce the amount of time available for therapy. In the future, we will thus also try to simplify the sensor setup by, e.g., determining which sensors can be removed without greatly affecting the overall affect classification accuracy.

B. Study Weaknesses

Aside from the lack of clinical evaluation, two study weaknesses should be acknowledged. First, since participants in a pair play the same game with the same hardware, the averaged responses over all "first" participants and all "second" participants should be very similar. However, this is not the case: for example, the two participants have different self-reported affects in some conditions (Table I), and classification accuracies are often different for the two participants (Table IV). This may be due to statistical noise in the relatively small sample, but may be due to unidentified systematic differences in the experiment setup.

Second, the classification design was not optimal. While we successfully induced different affects in the 6 conditions (Table I), classification of the 4 affects into 'low' and 'high' classes may be problematic since participants may not accurately self-report their affective states and since it is difficult to determine exactly what "low" and "high" should correspond to on a Likert-type scale. In the future, we will consider alternative pattern recognition schemes, such as classifying balanced vs. unbalanced exercise or using regression to estimate the four affects on the same 7-point scale used by the short questionnaire.

V. Conclusion

We developed a method that automatically classifies players' affective states in a competitive arm rehabilitation game based on their physiological responses. The method incorporates analysis of each individual player's physiological responses as well as analysis of the covariation of the players' physiological responses (physiological linkage). The method resulted in good affect classification accuracies (75% to 91% in two-class classification), and physiological linkage provided important information that could not be obtained by analyzing each individual player's physiological responses. In the future, our classifiers will be combined with decision-making rules that will dynamically adapt different game difficulty settings based on both players' affective states, thus resulting in enjoyable, intensive rehabilitation for both players. These classifiers and decision-making rules will be tested with actual patients in both competitive and cooperative rehabilitation games, and the method may generalize to other human-computer interaction scenarios such as collaborative learning.

References

[1] K. E. Laver, B. Lange, S. George, J. E. Deutsch, G. Saposnik, and M. Crotty, "Virtual reality for stroke rehabilitation," Cochrane Database of Systematic Reviews. 2017.

[2] S. Lim and B. Reeves, "Computer agents versus avatars: Responses to interactive game characters controlled by a computer or other player," Int. J. Hum. Comput. Stud., vol. 68, no. 1-2, pp. 57-68, Jan. - Feb. 2010.

[3] D. Weibel, B. Wissmath, S. Habegger, Y. Steiner, and R. Groner, "Playing online games against computer- vs. human-controlled opponents: Effects on presence, flow, and enjoyment," Comput. Human Behav., vol. 25, no. 5, pp. 2274-2291, Sep. 2008.

[4] M. Mace, N. Kinany, P. Rinne, A. Rayner, P. Bentley, and E. Burdet, "Balancing the playing field: Collaborative gaming for physical training," J. Neuroeng. Rehabil., vol. 14, article no. 116, Nov. 2017.

[5] M. Goršič, I. Cikajlo, and D. Novak, "Competitive and cooperative arm rehabilitation games played by a patient and unimpaired person: effects on motivation and exercise intensity," J. Neuroeng. Rehabil., vol. 14, article no. 23, Dec. 2017.

[6] M. Goršič, I. Cikajlo, N. Goljar, and D. Novak, "A multisession evaluation of an adaptive competitive arm rehabilitation game," J. Neuroeng. Rehabil., vol. 14, article no. 128, 2017.

[7] K. M. Triandafilou, D. Tsoupikova, A. J. Barry, K. N. Thielbar, N. Stoykov, and D. G. Kamper, "Development of a 3D, networked multi-user virtual reality environment for home therapy after stroke," J. Neuroeng. Rehabil., vol. 15, article no. 88, Oct. 2018.

[8] K. Baur et al., "Robot-supported multiplayer rehabilitation: Feasibility study of haptically linked patient-spouse training," in 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2018, pp. 4679–4684.

[9] N. Norouzi-Gheidari, M. F. Levin, J. Fung, and P. Archambault, "Interactive virtual reality game-based rehabilitation for stroke patients," in 2013 International Conference on Virtual Rehabilitation (ICVR), 2013, pp. 220–221.

[10] M. Pirovano, R. Mainetti, G. Baud-Bovy, P. L. Lanzi, and N. A. Borghese, "Intelligent game engine for rehabilitation (IGER)," IEEE Trans. Comput. Intell. AI Games, vol. 8, no. 1, pp. 43–55, Mar. 2016.

[11] D. Novak, M. Mihelj, J. Ziherl, A. Olenšek, and M. Munih, "Psychophysiological measurements in a biocooperative feedback loop for upper extremity rehabilitation," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 19, no. 4., pp. 400-410, Aug. 2011.

[12] C. Rodriguez-Guerrero, K. Knaepen, J. C. Fraile-Marinero, J. Perez-Turiel, V. Gonzalez-de-Garibay, and D. Lefeber, "Improving challenge/skill ratio in a multimodal interface by simultaneously adapting game difficulty and haptic assistance through psychophysiological and performance feedback," Front. Neurosci., vol. 11, no. 242, May 2017.

[13] A. Darzi, M. Goršič, and D. Novak, "Difficulty adaptation in a competitive arm rehabilitation game using real-time control of arm electromyogram and respiration," in IEEE International Conference on Rehabilitation Robotics, 2017.

[14] M. E. Vanutelli, L. Gatti, L. Angioletti, and M. Balconi, "Affective synchrony and autonomic coupling during cooperation: a hyperscanning study," Biomed Res. Int., vol. 2017, no. 3104564, 2017.

[15] Y. Hu, Y. Pan, X. Shi, Q. Cai, X. Li, and X. Cheng, "Inter-brain synchrony and cooperation context in interactive decision making," Biol. Psychol., vol. 133, pp. 54-62, Mar. 2018.

[16] C. Szymanski et al., "Teams on the same wavelength perform better: Inter-brain phase synchronization constitutes a neural substrate for social facilitation," Neuroimage, vol. 152, pp. 425-436, May 2017.

[17] M. M. Spapé, J. M. Kivikangas, S. Järvelä, I. Kosunen, G. Jacucci, and N. Ravaja, "Keep your opponents close: Social context affects EEG and fEMG linkage in a turn-based computer game," PLoS One, vol. 8, no. 11, pp. e78795, 2013.

[18] N. Sinha, T. Maszczyk, W. Zhang, J. Tan, and J. Dauwels, "EEG hyperscanning study of inter-brain synchrony during cooperative and competitive interaction," in 2016 IEEE International Conference on Systems, Man, and Cybernetics, 2016.

[19] G. Chanel et al., "Grand Challenge Problem 2: Adaptive Awareness for Social Regulation of Emotions in Online Collaborative Learning Environments," in Grand Challenge Problems in Technology-Enhanced Learning II: MOOCs and Beyond, 2016, pp. 13–16.

[20] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," Front. Public Heal., vol. 5, no. 258, Sep. 2017.

[21] W. Boucsein, "Electrodermal activity," in Methods of Electrodermal Recording, Boston, MA: Springer US, 2012, pp. 87–259.

[22] I. Ekman, G. Chanel, S. Järvelä, J. M. Kivikangas, M. Salminen, and N. Ravaja, "Social interaction in games," Simul. Gaming, vol. 43, no. 3, pp. 321–338, Jun. 2012.

[23] S. Järvelä, J. M. Kivikangas, J. Kätsyri, and N. Ravaja, "Physiological linkage of dyadic gaming experience," Simul. Gaming, vol. 45, no. 1, pp. 24–40, Feb. 2014.

[24] S. M. Alessio, Digital Signal Processing and Spectral Analysis for Scientists. Cham: Springer International Publishing, 2016.

[25] Keough & Quinn, "Multiple regression and correlation," Des. Anal. Biol., 1995.

[26] S. McCrea, G. Geršak and D. Novak, "Absolute and relative user perception of classification accuracy in an affective video game," Interact. Comput., vol. 29, no. 2, pp. 271-286, Mar. 2017.

[27] N. Goljar et al., "Psychophysiological responses to robot training in different recovery phases after stroke," in 2011 IEEE International Conference on Rehabilitation Robotics, 2011.