Neurogaming-based Classification of Player Experience Using Consumer-Grade Electroencephalography

Thomas D. Parsons
Timothy McMahan
Ian Parberry

Technical Report LARC-2015-04

Laboratory for Recreational Computing
Department of Computer Science & Engineering
University of North Texas
Denton, Texas, USA

April, 2015

UNIVERSITY OF NORTH TEXAS

Neurogaming-based Classification of Player Experience Using Consumer-Grade Electroencephalography

Thomas D. Parsons, Timothy McMahan, and Ian Parberry

***Abstract*—A growing body of literature has emerged that demonstrates the potential of neurogaming platforms for interfacing with well-known video games. With the recent convergence of advances in consumer electronics, ubiquitous computing, and wearable sensor technologies real-time monitoring of neurocognitive and affective states can be studied in an objective, timely, and ecologically valid manner. Whilst establishing the optimal relation among frequency bands, task engagement, and arousal states is one of the main goals of neurogaming, a standardized method has yet to be established. Herein we aimed to test classifiers within the same context, group of participants, feature extraction methods, and protocol. Given the emphasis upon neurogaming, the commercial-grade Emotiv EPOC headset was used to collect electroencephalographic (EEG) signals from users as participants experienced various stimulus modalities aimed at assessing cognitive and affective processing. The EEG data were then filtered to get separate frequency bands to train cognitive-affective classifiers with three classification techniques: Support Vector Machines (SVM), Naive Bayes (NB), and k-Nearest Neighbors (kNN). Results revealed that the NB classifier was the most robust classifier for identifying game-based Death Events. However, the identification of General Gameplay Events is best identified using kNN and the Beta band. From this study's findings, it is suggested that using a combination of classifiers is preferable over selecting just one classifier.**

***Index Terms*—Neurogaming, Electroencephalography, Emotiv, Cognitive, Affective, Engagement, Arousal.**

I. Introduction

NEUROGAMING platforms use off-the-shelf brain computer interfaces (BCIs) to improve gameplay. Using a neurogaming electroencephalography (EEG)-based BCI platform, gamers can interact with a console without the use of a traditional controller. The addition of neurogaming to human-computer interaction is growing in popularity and the traditional use of a keyboard and mouse for gameplay is lessened as more natural touch and gesture interfaces become a more widely used interaction modality. The BCI technology used in neurogaming is growing rapidly and adoption of this technology is increasing. Although BCIs were traditionally developed as a communication tools, they have also been used for near real-time decoding of a person's neurocognitive or affective state. With the advent of neurogaming, researchers are increasingly incorporating BCIs into games. Neurogaming platforms follow neuroergonomic principles, in which knowledge of brain-behavior relationships can be applied to optimize the design of gaming environments to accelerate user experience. With the recent convergence of advances in consumer electronics, ubiquitous computing, and wearable sensor technologies real-time monitoring of neurocognitive and affective states can be studied in an objective, timely, and ecologically valid manner.

Video games represent an immersive activity that is rapidly increasing in popularity. According to the Entertainment Software Association [1], 72% of the general population and 97% of teenagers reported regular playing of video games. Further, video games are played more frequently and in more locations [2]. A growing body of literature has emerged that focuses upon cognitive assessment of video gamers [3]. In addition to the growth of video games in general, recent studies have demonstrated the potential of neurogaming applications for interfacing with well-known games as "Pacman" [4], "Tetris" [5], and "World of Warcraft" [6]. Given the growing popularity of neurogaming and the increasing literature on cognitive and affective aspects of video gamers, there is a need for novel approaches to assessment of cognitive and affective processes occurring while players are immersed in video games.

A. Knowledge of User-State During Video Gameplay

In addition to acting as a control interface, the EEG-based BCIs used in neurogaming can provide knowledge of user-state during video gameplay. This information is imperative for development and assessment of video game design [7]. Individuals will invariably have different reactions to a given game, and without an assessment tool that can be employed online, researchers will experience difficulties in identifying the causes of these differences. The EEG data from BCIs provide signals that are continuously available [8] and logged without the gamer's conscious awareness. This creates an objective measure of the gamer's state, which can include measures of cognitive workload [9], stress levels [10], task engagement [11], among others. EEG-based arousal and engagement indices can be gleaned from various sensors continuously, which further increases experimental control [12].

B. EEG for Establishing Indices of Engagement

Using EEG to measure task engagement is not a new concept. It has been widely used with medical grade EEG

T. Parsons is with the Department of Psychology, University of North Texas, Denton, TX, 76203.
E-mail: Thomas.Parsons@unt.edu

T. McMahan and I. Parberry are with the Department of Computer Science & Engineering, University of North Texas, Denton, TX, 76203. E-mail: Fred.McMahan@unt.edu, ian@unt.edu

Manuscript received April 1, 2015; revised April 1, 2015.

devices. Pope et al. [11] built a system to control the level of task automation based upon the whether the operator had increasing or decreasing engagement. Freeman et al. [13] expanded on this same system by evaluating performance of each task along with using absolute values of engagement versus just looking at increasing and decreasing engagement. Task engagement and mental workload are areas that Berka et al. [9] explored as a way to help identify more accurate and efficient methods for people to interact with technology with the possibility of developing more efficient work environments that increase motivation and productivity. Their results suggest that EEG engagement reflects information gathering, visual processing, and attention allocation. Smith and Gevins [14] used a flight simulator to subject participants to low, moderate or high difficulty tasks to see how the brain responded. The results from their study showed an increased frontal theta response along with reduced parietal alpha during demanding tasks. Other work [15], [16] had similar results indicating that an increase in theta and a decrease in alpha was correlated with increased number of tasks along with amount time a persons is awake. Yamada [17] measured frontal theta activity along with eye blinking and found that children playing video games had higher theta activity along with a high degree of blink inhibition. These results suggest that interesting tasks result in higher frontal theta activity while the task inhibits eye blink activity. Recently, Kamzanova et al. [18] compared the sensitivity of various EEG engagement indices during time-on-task effect and cueing to detect which index was most effective for detecting reduced alertness linked with vigilance decline in performance.

C. EEG to Isolate Specific Game Events

In the frequency domain, the spectral power in various frequency bands has been used for assessing arousal and affective states [19]. Beta, EEG coherence has been found to increase when participants viewed highly arousing stimuli [20]. Theta power event-related synchronization studies have found modulation during transitions in affective state [21]. In addition to spectral power and waveforms, interactions between pairs of EEG oscillations – such as phase synchronization and coherence – have also been implicated in affective states of hedonic arousal [22]. It has been suggested that higher frequency bands may have greater contribution contribute to arousal response than lower frequency bands [23]. Often, researchers emphasize the potential of Alpha power variance with the negative and positive valence states [24] or with discrete affective states such as happiness, sadness, and fear [25]. Alpha power frontal asymmetry has been repeatedly reported as a steady correlate of valence [26]. Subsequent findings have suggested that frontal Alpha asymmetry may reflect the approach/avoidance aspects of emotion [27]. The Gamma band has been shown in previous research to find changes in affect. Further, Gamma power event-related synchronization and desynchronization has been related to affective states such as happiness and sadness [28]. Also, increases in the gamma phase synchronization index have been induced by unpleasant visual stimuli [29].

D. Neurogaming Using Emotiv EPOC EEG

One of the most widely studied of the inexpensive off-the-shelf neurogaming systems is the Emotiv EPOC. It is a compact, wireless headset that requires comparatively little effort to set up. It allows increased flexibility and mobility over traditional EEG. Thus, providing an inexpensive tool that game developers can use to measure EEG. Although the Emotiv is aimed at the gaming market and is not classified as a medical device, researchers have adopted it for a variety of applications [30], [31]. Using the Emotiv, researchers can detect facial movements, emotional states, and imagined motor movement. Although the Emotiv EEG does not have the fidelity of a laboratory EEG it still offers the ability to provide a gamer's brainwave signature. The system has been found to work well detecting focused thoughts [32], [33]. Duvinage et al. [34] compared the Emotiv headset to the Advance Neuro Technology (ANT) acquisition system during a run with the P300 speller system. Although the Emotiv headset was not found to be as accurate as the ANT system (a medical grade device), it was able to capture EEG signal at a successful level that was deemed adequate for games.

Researchers have also investigated different EEG processing algorithms to assess classification of shapes being thought about [35], detection of hand movement intentions on the same side of the brain as the hand [36], classification of positive and negative emotion elicited by pictures [37], and evaluation of cognitive workload [38]. With the benefit of being noninvasive to the wearer, it is a tool that is practical for use by game developers. McMahan et al. [39] were able to find significant difference in the Beta and Gamma bands among various stimulus modalities. They also found an increase in the power estimates during high intensity game play (e.g., death events) when compared to low intensity general game play. The authors conclude that their findings suggest that the Emotiv EEG can be used to assess differences in frequency bands when persons are experiencing various stimulus modalities using off-the-shelf EEG-based gaming technology. In addition to task engagement, affective states have been measured while users watched a film [40].

In sum, various neurogaming platforms use a gamer's psychophysiological indices to complete tasks or affect the mood of the game. To date, the research design, data logging of game-based psychophysiological signals, and the control algorithms found in neurogaming are not systematic and studies to support their use remains limited. Further, psychological studies regarding the relations between affective and cognitive correlates of brain processing are uncovering the strong implication of cognitive processes in emotions [41]. This has resulted in increasing emphases upon affective neuroscience [42] and the potential for EEG data to proffer valuable information about the participants' felt cognitive and affective processing. Although there have been growing efforts in the neurogaming literature to recognize a user's cognitive and affective states in real time using EEG, these indices are typically developed in isolation and do little to take into account bout cognitive and affective information. While establishing the optimal relation among frequency bands, task engagement, and arousal states is one of the main goals of neurogaming, a standardized method has yet to be established. The ideal research situation would test classifiers within the same context, users, feature extraction methods, and protocol [43].

As neurogaming systems increase in use, new properties will need to be taken into consideration. A common difficulty encountered in this research area is the dearth of published objective comparisons among classifiers. In this paper, an approach to task engagement and affective state estimation for neurogaming is explored. The EEG signals from users were logged as participants experienced various stimulus modalities aimed at assessing cognitive and affective processing. Given the emphasis upon neurogaming, the commercial Emotiv EPOC headset was used. The EEG data were then filtered to get separate frequency bands to train cognitive-affective classifiers with three classification techniques: Support Vector Machines, Naive Bayes, and k-Nearest Neighbors. The current study aimed to develop arousal and engagement indices to assess various levels of gaming experience using the Emotiv EEG.

II. Methods

A. Participants

Electroencephalographic (EEG) data was collected from 30 college aged students (66% female, mean age = 20.87, range 18 to 43). Participants were recruited from university undergraduate and graduate pools; education levels ranged from 13 to 20 years. Ethnicity was as follows: Caucasian (n=20), African American (n=1), Hispanic (n=4), Native American (n=1), and Asian Pacific (n=4). All participants were right handed and had at least average computer skills. Game playing skills ranged from casual cell phone games to playing every day on a personal computer or a game console. All participants endorsed use of a computer at least once per day (every day) with 30% reporting use of a computer several times a day.

In terms of computer experience, 66% rated themselves as experienced, 27% rated themselves as somewhat experienced, and 7% rated themselves as very experienced when ranking their computer competency. Homogeneity of the sample was found in that there were no significant differences among participants relative to age, education, ethnicity, sex. Further, there were no significant differences relative to, self-reported symptoms of depression, sleepiness, game play experience, or computer use. Strict exclusion criteria were enforced to minimize possible confounding effects of comorbid factors known to adversely impact cognition, including psychiatric conditions (e.g., mental retardation, psychotic disorders, diagnosed learning disabilities, attention deficit/hyperactivity disorder, and bipolar disorders, as well as substance-related disorders within 2 years of evaluation) and neurologic conditions (e.g., seizure disorders, closed head injuries with loss of consciousness greater than 15 minutes, and neoplastic diseases). The participants received class credit for their participation in the study.

B. Apparatus and Measures

1) Super Meat Boy: Super Meat Boy [44] is a platform game in which players control a small, dark red, cube-shaped character named Meat Boy. The participant played a cube of meat jumping around the level to avoid saw blades to reach their goal of rescuing bandage girl. This game requires the minimum amount of keys to play (arrow keys and space bar) thus making it easy for any level of gamer to achieve success. Major events in the game include successfully completing a level and dying. Dying occurs from running into spinning saw blades or falling into fire. As the player progresses through the game the levels get increasingly difficult by adding more saw blades and large jumps. Each level is timed as a goal of the game is to get through each level as fast as possible. The core gameplay requires fine control and split-second timing [45]. Primary game events used for this study included: 1) Death events; and 2) “General Game Play”. The “Death events” occurred when the participant’s character died. Although there are a number of possible ways for a character to die in a game (e.g., the character gets sliced to pieces, or falls into acid, or gets skewered on needles), we sampled from death events related to the character falling into acid. The “General Game Play” was differentiated from “Death events” in that general game play was sampled during periods in which the player had not experienced any death events for 1 minute before or after “General Game Play” sampling.

2) Two-Picture Cognitive Discrimination Task: Participants were shown a pair of color pictures of a landscape, and were given the evaluative task of identifying any differences between the pair. Unknown to the participants, the pictures were identical.

3) Spider Jump Arousal Stimulus: The Spider Jump Arousal stimulus was first storyboarded and designed on paper. A 3-D model of a venomous headcrab was taken from the Half-Life 2 game [46]. The venomous headcrab was chosen because it leaps with incredible speed while releasing an angry squeal when a suitable host is in a clear line of sight. The participants were subjected to the Spider Jump Arousal Stimuli without any cue or knowledge that it would occur.

4) Game Experience Survey: Participants answered a series of questions assessing their prior video game experience and other personal characteristics. Participants were asked to report the number of hours they spent playing video games on their cell phones (M = 3.47), playing games on their computer (M = 3.47), and playing games on their game console (M = 2.3). 20% of the participants reported playing video games more than 20 hours per week. The participants were also asked if they would classify themselves as “gamers”, 33% responded as being part of this category.

5) Emotiv EPOC EEG: This Emotiv EEG headset has 14 electrodes (saline sensors) locating at AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1, O2 (see Fig. 1) and two additional sensors that serve as CMS/DRL reference channels (one for the left and the other for the right hemisphere of the head). The Emotiv EEG’s 14 data channels are spatially organized using the International 10–20 system. The Emotiv EPOC headset does not require a moistened cap to improve

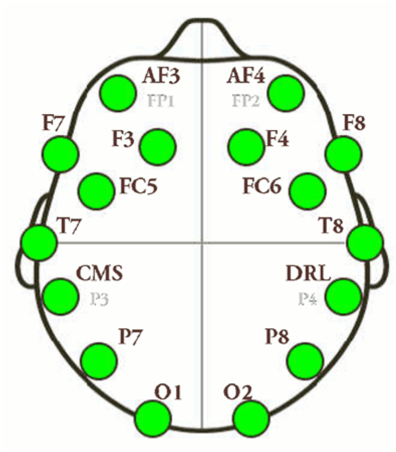

Fig. 1. Sensor locations on Emotiv headset.

conduction. The sampling rate is 128Hz, the bandwidth is 0.2-45Hz, and the digital notch filters are at 50Hz and 60Hz.

C. Procedure

Upon arriving at the testing office, the participants were given an informed consent to read and sign. Included in the informed consent was a waiver to record the participant during the study. The participants were then seated in a comfortable chair and given a keyboard and mouse to complete a questionnaire about computer and game experience. For the actual assessment, each participant played the game in the same room location. The game was displayed on a Samsung 60 inch plasma screen. The participants sat in a chair that has a built in keyboard tray, along with a speaker system and USB port around head level to minimize the distance between the Emotiv headset and the receiver/transmitter. While the participant played the game the lights were turned off to help immerse the player into the game and reduce glare from the overhead lights. The experimenter combed the participants on the left, mid-line, and right sides of their scalp firmly in order to reduce electrode impedances [47]. After the relevant areas on the face and mastoids had been cleaned, the Emotiv EEG headset was positioned on the participant's head. The examiner verified impedances in connections between each electrode and the participant's scalp.

To establish base line for each participant a video was played and were told "to relax and try not to think about anything". In the video the screen was blank for 2:00 minutes to establish a minimum brain wave activity. Next, the participant completed two tasks: the Two Picture Cognitive Task and Super Meat Boy task (see Fig. 2). Task presentation order was counter-balanced across participants. After the initial task (either Two Picture Cognitive Task or Super Meat Boy) the participant was presented with 1:30 seconds of blank screen viewing to allow the participant to return to a steady state. During the Two-Picture Cognitive Task they compared two pictures to determine the difference between them. This allowed for the establishment of a brain wave signature for basic cognitive processing. During the Super Meat Boy Task the researcher aided the participant with the first few levels to allow the player to become acquainted with the rules and game controls. Next, participants were informed that they would play Super Meat Boy for 15 minutes and that they were to advance as far as they could in the game.

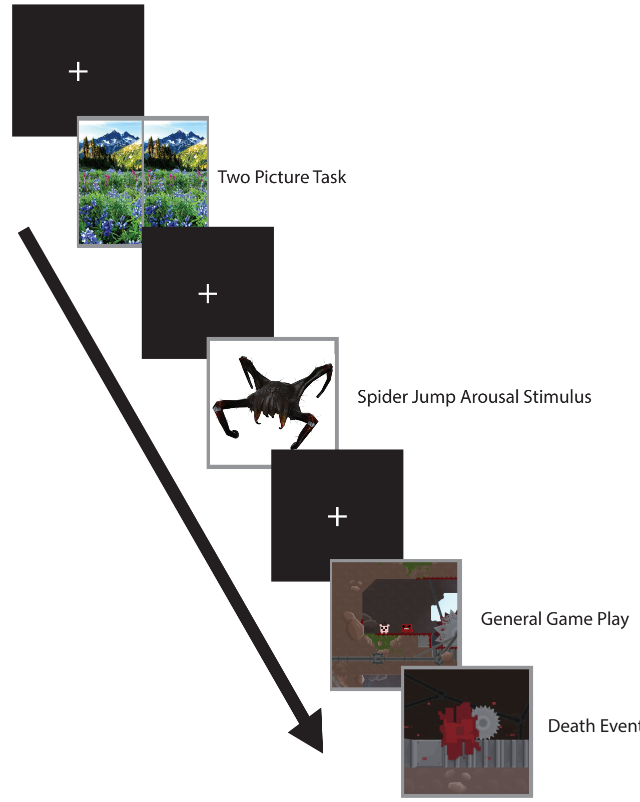

Fig. 2. Chronological order of events that participants encountered.

Each participant's game play was captured in 1080p HD (60 frames per second) using a Hauppauge video capturing device allowing the game play to be synced the EEG data. Each participant was also recorded using a Logitech 9000 HD webcam to help isolate events (facial or body movements) that may affect the EEG data. EEG data and video data were recorded on the same computer with all non-essential programs closed. Using OpenViBE drift correction, a 128 Hz sample rate was achieved minimizing any syncing issues between the EEG data and the video recording of game play. Syncing all video recordings with EEG recording software involved the use of screen captures before and after every section of the study (baseline video and game play). Each screen shot produced a time stamp for EEG data and video to establish the location of the start and end of each section. The screen shots were saved to reference later during the data analysis phase.

D. Data Analytics

All data were analyzed using SAS version 9.1. Descriptive statistics were calculated for participant demographics and for EEG results. Missing data were imputed by either mean substitution or last case carried forward.

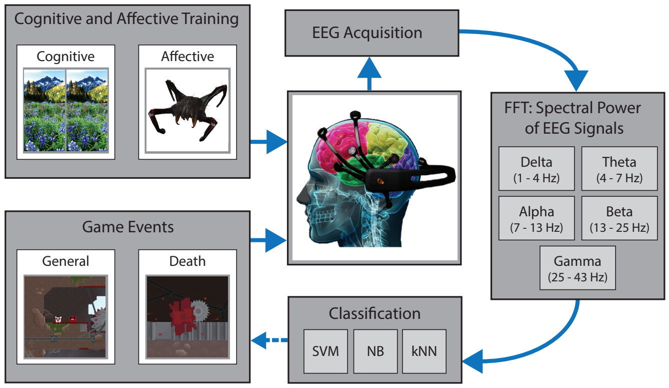

Fig. 3. Flow chart depicting the procedure used to obtain results. The dashed line between classification and game events signifies the eventual use of a closed loop system.

The Emotiv Epoc headset was used to capture the EEG data from each participant (see Fig. 3). Emotiv TestBench and OpenViBE were used to capture the raw EEG output from the headset. The EEG data was segmented into epochs that started 100 ms before the onset of each stimulus (0 ms), and ended 750 ms after the onset of the same stimulus. Epochs were calculated for 4 different modalities: 1) Two-Picture Cognitive Task; 2) Spider Jump Arousal Stimulus; 3) General Game Play; and 4) Death events.

Artifacts such as blinking, head movements, or body movement can cause unwanted data in EEG data. Most EEG analysis requires removal of these artifacts to help identify medical issues. However this is not necessarily a detrimental issue when using for game play analysis. These types of artifacts are common in everyday game play [48]. These artifacts can actually be used for further analysis as body movement or other movement can signify engagement [49]. The EEG artifact data was annotated as artifact where visually noticeable deflection in the EEG was observed at the times that participants performed movements. Artifacts related to eye blinks and other muscle movements in addition to physical movements of the sensors themselves were removed before the EEG traces were processed. The Emotiv SDK automatically detects and records eye blinks. Given that muscle contraction and control are generally governed outside of the frequency range of interest [50], we were able to use frequency band limiting procedures such as low-pass, high-pass and notch filters to adequately remove these signal components. As Anderson et al. [51] describe, after removing EEG artifacts the researcher may assess whether the energy densities of the alpha or theta frequency bands are changed by more than 20% of their original values. If so, the trial should be removed from all further analysis. In this study, we did not need to throw out any of the trials due to excessive signal degradation from movement or excessive change in spectral densities.

The spectral power of EEG signals in different bands has been found to be correlated with emotions [52]. The power estimates (μV^2) were found using a fast Fourier transform (FFT) and a 1 second Hamming window with no overlap for Delta (1 – 4 Hz), Theta (4- 7 Hz), Alpha (7 -13 Hz), Beta (13 – 25 Hz) and Gamma (25 – 43 Hz) for all 14 sensor location on the Emotiv headset. In typically EEG studies, the number of channels (e.g., 32, 64, 128, or 256 EEG channels) ranges from 32 channels (for routine exams) up to 256 channels (for source estimation) and the systems are able to sample at up to 1000Hz. Given that the Emotiv has only 14 channels and the data sample rate is only 128Hz, the average was calculated across all 14 sensors to obtain a global average for each frequency band. Following Anderson et al. [38] the baseline and stimulus signals were transformed to determine the power change and frequency shift induced by the task. These values are used to calculate the cognitive processing experienced at each of the 14 sensors for a given task. The spatial averaging of the 14 values gives a single measurement for analysis. Finally the data was normalized with the natural logarithm (ln).

Pope et al. [11] and Freeman et al. [13] have shown that an engagement index can be calculated by taking the ratio of Beta / (Alpha + Theta) EEG bands. Berka et al. [9] was able to show that the engagement index reflected a person's process of information-gathering, visual scanning and sustained attention. The engagement index was calculated for each participant using the single measurement form all sensors. Arousal has been shown to be measured by using (BetaF3 + BetaF4) / (AlphaF3 + AlphaF4) and valence using (AlphaF4 / BetaF4) - (AlphaF3 / BetaF3) [53].

All user data sets were analyzed together aiming to verify the possibility of building a generalizable model. We assessed the importance of all the EEG signals and their aggregate impact on the classification accuracy. Time epochs were split into corresponding signals, which resulted in 128 EEG measurements each for Alpha, Beta, Theta, Gamma, Engagement Index, and Arousal Index. This produced 30 data sets for each event: Two Picture Cognitive Discrimination Task – 30 sets of data to train representing one from each participant; Spider Jump Arousal Stimulus – 30 sets of data to train representing one from each participant; Game Play - 30 sets of data to test representing one from each participant; and Death Event - 30 sets of data to test representing one from each participant. Training was completed on the Two Picture Cognitive Discrimination Task and Spider Jump Arousal Stimulus data sets and tested on the Game Play and Death data sets. Given that it was a 50/50 data set, cross-validation was irrelevant and it was not used.

1) Support Vector Machine: To classify a set of binary labeled data, the support vector machine (SVM) algorithm uses a hyperplane to separate the data into two classes. During the training process of the SVM takes in data belonging to each category and maps them into a higher dimensional space with the goal of creating a hyperplane with the maximum difference. The training process can use different types of kernels (linear, polynomial, or radial basis function) to achieve a better hyperplane. During the testing process test new data is run through the SVM and placed into one of two categories based upon which side of the hyperplane the new point falls following training of the algorithm on a given data set, the discriminate hyperplane is optimized and selected based on the maximum margins between the hyperplane and the data. This is accomplished via transformation of the data from the input space into feature space (in which linear classification is achievable). This is achieved through outlier accommodating and error allowance during training [54]. The SVM technique has been used for arousal state estimation and results revealed a recognition accuracy of 83% could be achieved [55]. Herein the Classification SVM Type 2 was implemented in the libsvm library using 0.5 nu-SVM classification with radial basis function kernel. Gamma was set to 0.008 and the maximum number of iterations 1000. A stop error of 0.001 was utilized.

2) Naive Bayes: The Naive Bayes (NB) Classifier technique is based on Bayes theorem and is appropriate when the dimensionality of the inputs is high. This classifier computes the probability that some data points belong to a specific class. To perform the classification, the algorithm chooses the class with the highest probability, as its result. When event related potentials are included as a feature, the NB has been used to classify emotions in two classes (low valence and high valence) with a classifying accuracy of 56% [56]. In a related analysis, NB has been found to provide recognition accuracy of 70% for two classes (as reviewed in Nie et al. [23]). The NB is an efficient supervised learning algorithm used to classify data into different groups based upon a calculated probability of new data belonging to that group. The NB classifier makes the assumption that each input is independent from every other input. During the training phase the classifier takes the inputs and builds feature vectors for each category. When new data is presented to the NB classifier it uses the maximum likelihood estimates to find place that data into the correct category. The NB classifier has an added benefit of not requiring large sets of training data to be effective at classification.

3) k-Nearest Neighbor: k-Nearest Neighbor (kNN) is a supervised learning algorithm that classifies data into different groups based upon how closes it is located to a category. During training the classifier stores each category data into a feature vector. New data is then classified based upon the training sample that has the shortest distance to the new data point. An issue that can arise from the kNN classifier is if the data does not have an even distribution causing the classifier to favor one category over the other. In a study of arousal state estimation Lin et al. [57] extracted power spectrum density of different EEG sub-bands as features during an emotion induction (listening to music) protocol. They found a classification accuracy of 82% for four emotions. In another study using the kNN technique for two different sets of EEG channels (62 channels and 24 channels), an accuracy of 82.87% was found for the 62 channel data set and 78.57% for the 24 channel data set for five emotions [58].

TABLE I
OVERALL CLASSIFIER PERCENTAGES

Machine Learning	Mean	Std. Deviation	Min	Max
SVM	57.5	4.44	50.0	63.3
NB	70.0	6.56	58.6	75.9
kNN	57.9	11.15	44.5	76.0

TABLE II
SIGNAL CLASSIFICATIONS PERCENTAGES

Machine Learning	Mean	Std. Deviation	Min	Max
Engagement	64.1	7.50	56.7	71.6
Arousal	59.6	5.10	54.3	64.4
Alpha	55.4	14.34	44.5	71.6
Beta	68.9	10.64	56.7	76.0
Theta	61.4	13.20	50.0	75.9
Gamma	60.1	2.76	58.5	63.3

E. Results

Each participant's results from the Two Picture Cognitive Discrimination Task and Spider Jump Arousal Stimulus were used to predict General Gameplay Events and Death Events using a Support Vector Machine (SVM), a Naive Bayes (NB) classifier, and a k-Nearest Neighbor (kNN) classifier (see Table I). Having thirty participants in the study allowed for a total of 60 data points (30 for the Two Picture Cognitive Discrimination Task and 30 for the Spider Jump Arousal Stimulus) to train each classifier and 60 data points to test each classifier (30 for the General Gameplay Events and 30 for the Death Events). 128 EEG measurements were used for the predictors which represent the time period for each modality. The Engagement Index (Beta / (Alpha + Theta); Pope et al. [11] and Freeman et al. [13]), Arousal Index (BetaF3 + BetaF4) / (AlphaF3 + AlphaF4) and Valence Index (AlphaF4 / BetaF4) - (AlphaF3 / BetaF3; [53]), as well as Alpha, Beta, Theta, and Gamma bands were all tested to identify the strongest signals for classification (see Table II and Table III).

1) Machine Learning Classifiers and EEG Power Spectral Bands: Fig. 4 shows the overall accuracy for each classifier using the different signals. From Fig. 4 it is apparent that the strongest classifier was NB especially when using the Theta and Beta signals. The NB classifier had an overall average of 70% correct classification. Although the kNN classifier produced the highest accuracy rate with the Beta signal when compared to other classifiers, it performed poorly with the Alpha signal. Gamma turned out to be the strongest predictor in the SVM classifier. The Beta band wave was the strongest predictor followed by Theta, the Engagement Index, and then Alpha.

2) Distinguishing between General Game Play and Death Events: Fig. 5 illustrates that using the Two Picture Cognitive Discrimination Task and the Spider Jump Arousal Stimulus to train the SVM classifier did a better job overall classifying General Gameplay Events over Death Events. The strongest signals again were Beta, Theta, and the Engagement Index. The Gamma band showed the most potential with this classifier as it did the best job in classifying Death Events.

Fig. 6 shows that the NB classifier did the best job classifying Death Events from the training data. The strongest signals were Theta and Beta followed by Alpha and the

TABLE III
INDIVIDUAL CLASSIFICATION PERCENTAGES

Signal	SVM	NB	kNN
Engagement	56.7	71.6	64.0
Arousal	60.0	64.4	54.3
Alpha	50.0	71.6	36.1
Beta	56.7	74.0	76.0
Theta	58.3	75.9	50.0
Gamma	63.3	58.6	58.5

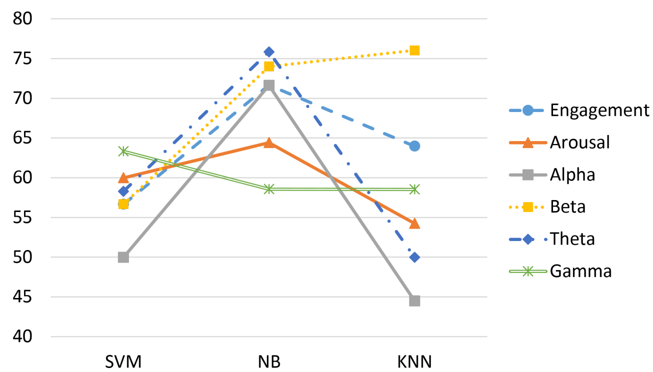

Fig. 4. Overall classifier results for each machine learning algorithm.

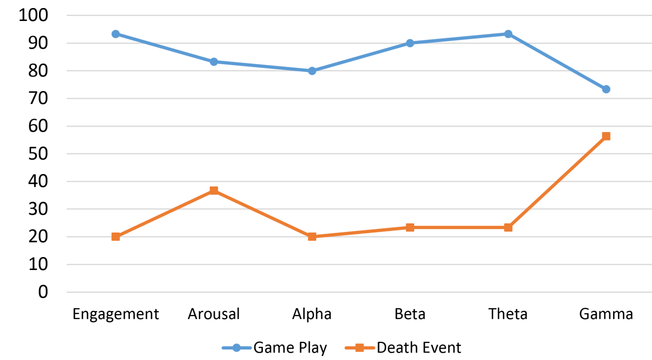

Fig. 5. SVM classifications percentage across each signal for game play and death events.

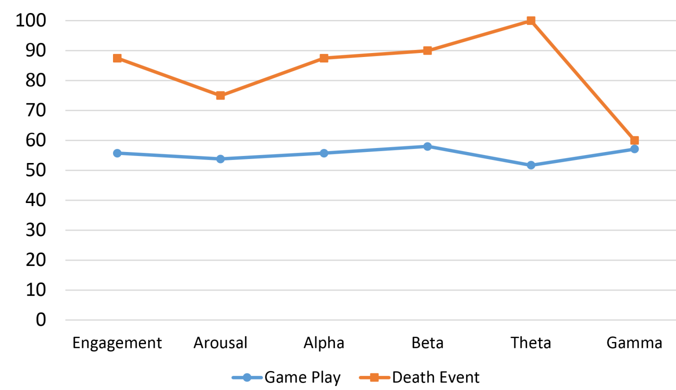

Fig. 6. NB classification percentages across each signal for game play and death events.

Engagement Index. While Gamma performed well for the SVM classifier, the Gamma band performed the worst for NB. The overall trend of the NB classifier reveals it as being the most steady and reliable in distinguishing between General Gameplay Events and Death Events.

Although the kNN classifier had the greatest variance in terms of signal was being used for classification, it did a better job overall with General Gameplay Events (see Fig. 7). The Beta signal was the strongest predictor for the kNN classifier for both General Gameplay Events and Death Events. Although Alpha was the weakest predictor, it performed better in predicting Death Events then General Gameplay Events. The overall trend of the kNN classifier was erratic but revealed potential when using the Beta signal.

III. DISCUSSION

While various neurogaming platforms use machine learning to model a gamer's EEG indices, the research designs, data logging of game-based psychophysiological signals, and the control algorithms found in neurogaming are not systematic and studies to support their use remains limited. As neurogaming systems increase in use, new properties will need to be taken into consideration. A common difficulty encountered in this research area is the dearth of published objective comparisons among classifiers. Although there have been growing efforts in the neurogaming literature to recognize a user's cognitive and affective states in real time using EEG bands, these studies do little to take into account both cognitive and affective information. While establishing the optimal relation among frequency bands, task engagement, and arousal states is one of the main goals of neurogaming, a standardized method has yet to be established. Herein we aimed to test classifiers within the same context, users, feature extraction methods, and protocol [43]. Specifically, the EEG signals from users were logged as participants experienced various stimulus modalities aimed at assessing cognitive and affective processing. Given the emphasis upon neurogaming, the commercial Emotiv EPOC headset was used. The EEG data were then filtered to get separate frequency bands to train cognitive-affective classifiers with three classification techniques: Support Vector Machines, Naive Bayes, and k-Nearest Neighbors.

A. Machine Learning Classifiers and EEG Power Spectral Bands

The Beta band wave was the strongest predictor followed by Theta, the Engagement Index, and then Alpha. This was not surprising given that Beta EEG coherence has been found to increase when participants viewed highly arousing stimuli [20]. Further, McMahan et al. [39] found significant difference in the Beta band among various stimulus modalities. The Naive Bayes classifier had an overall average of 70% correct classification. Further, NB was found to be the strongest classifier when using the Theta and Beta signals. These findings are consistent with findings that NB has been found to have a good classification for two classes [56]. Although the kNN classifier produced the highest accuracy rate with the Beta

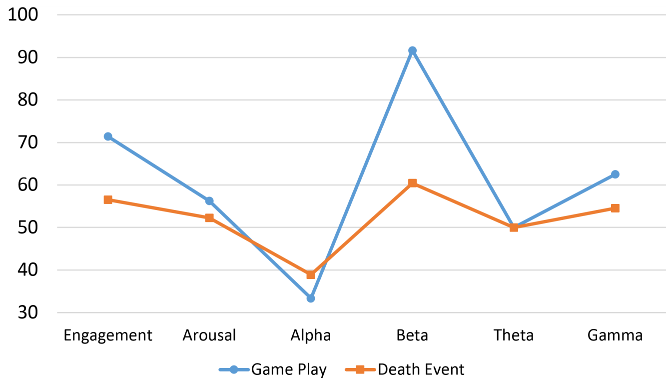

Fig. 7. kNN classification percentages across each signal for game play and death events.

signal than any other classifier, it performed poorly with the Alpha signal. For the SVM classifier, Gamma turned out to be the strongest predictor.

B. *Cognitive and Affective Training*

Using the Two Picture Cognitive Discrimination Task and the Spider Jump Arousal Stimulus to train the SVM classifier did a better job overall classifying General Game Play over Death Events. These results support findings that the SVM technique is useful for classifying arousal state and has been found to have a recognition accuracy of 83% [55]. Again, the strongest signals were Beta, Theta, and the Engagement index. The Gamma band showed the most potential with the SVM classifier as it did the best job in classifying Death Events. The Gamma band has been shown in previous research to find changes in emotion [59], however this research looked at a Gamma band ranging from 30-100 Hz which was far outside the range of the Emotiv (has a cut off of 45 Hz).

C. *Distinguishing between General Game Play and Death Events*

The NB classifier did the best job classifying Death Event from the training data. The strongest signals were Theta and Beta followed by Alpha and the Engagement index. Unlike in the SVM classifier the Gamma band performed the worst. The overall trend of the NB classifier shows it being the most steady and reliable in distinguishing between General Gameplay Events and Death Events. The kNN classifier varied more upon which signal was being used to classify, but overall did a better job with General Gameplay Events. The Beta signal was the strongest predictor in for the kNN classifier for both General Gameplay Events and Death Events. The Alpha signal was the weakest predictor, however it did perform better in predicting Death Events than General Gameplay Events. Although the overall trend of the kNN classifier was erratic, potential was observed when using the Beta signal.

D. *Real-time implementation*

This study provided an initial validation of an approach to using baseline EEG measures to predict various events users experience while playing video games. Given the validation resulting from the current study, future work will aim at developing a neurogaming protocol that includes training of classifiers off-line using baseline tasks so that during the subsequent game play events encountered by users could be readily identified. The results suggest that the NB classifier is the most robust classifier for identifying Death Events. However, the identification of General Gameplay Events is best identified using kNN and the Beta band. From this study's findings, it is suggested that using a combination of classifiers is preferable over selecting just one classifier. For example, a weighting scheme could be implemented in which weights are applied to the strongest attributes from each classifier (e.g., Beta signal from kNN). This would result in a more robust and powerful final determinate of which category best fits an existing user state. The results from all of the classifiers would then be compared to find the optimal state of the player. Whilst using multiple classifiers requires greater training time performed off-line, once trained and weighted, these classifiers will be executed during real-time gameplay. As events are successfully identified the classifiers can use those events to retrain the classifier to improve the accuracy.

E. *Limitations and Future Directions*

Our findings should be understood in the context of some limitations. These findings are based on a fairly small sample size. As a necessary next step, the reliability and validity of the Emotiv EEG needs to be established using a larger sample of participants to ensure that the current findings are not an anomaly due to sample size. Further, findings need further validation through straightforward comparison of Emotiv EEG results with those of standard laboratory-based EEG assessment technology. It is important to note, however, that the Emotiv has been favorably compared to a laboratory-based research EEG system (Neuroscan). Badcock et al. [60] found that the Emotiv EEG system can prove a valid alternative to laboratory ERP systems for recording reliable late auditory ERPs over the frontal cortices. While we found some interesting results, it is important to emphasize that these are very preliminary there are not currently well-established methodologies for examining the impact of game levels on players. Nevertheless, there is an increasing body of literature suggesting that game impact can be measured via EEG [61], [62]. Future studies will be needed to expand these results into methodological approaches to quantifying video game based EEG assessment in general and Emotiv–based EEG assessment of various games in particular.

IV. Conclusion

We have presented findings from a neurogaming protocol study aimed at using an off-the-shelf Emotiv EEG to test classifiers within the same context, users, feature extraction methods, and protocol. Results provided initial validation of an approach to using baseline EEG measures to predict various events users experience while playing video games. Given the validation resulting from the current study, future work will aim at developing a neurogaming protocol that includes training of classifiers off-line using baseline tasks so that during the subsequent game play events encountered by users could be readily identified.

References

[1] "Essential facts about the computer and video game industry: Sales, demographic, and usage data," 2010.

[2] V. J. Rideout, U. G. Foehr, and D. F. Roberts, "Generation m [superscript 2]: media in the lives of 8-to 18-year-olds." *Henry J. Kaiser Family Foundation*, 2010.

[3] M. S. Cain, A. N. Landau, and A. P. Shimamura, "Action video game experience reduces the cost of switching tasks," *Attention, Perception, & Psychophysics*, vol. 74, no. 4, pp. 641–647, 2012.

[4] B. Reuderink, A. Nijholt, and M. Poel, "Affective Pacman: A frustrating game for brain-computer interface experiments," in *Intelligent Technologies for Interactive Entertainment*. Springer, 2009, pp. 221–227.

[5] G. Pires, M. Torres, N. Casaleiro, U. Nunes, and M. Castelo-Branco, "Playing Tetris with non-invasive BCI," in *2011 IEEE 1st International Conference on Serious Games and Applications for Health*. IEEE, 2011, pp. 1–6.

[6] B. Van de Laar, H. Gurkok, D. Plass-Oude Bos, M. Poel, and A. Nijholt, "Experiencing BCI control in a popular computer game," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 5, no. 2, pp. 176–184, 2013.

[7] D. A. Norman, *Emotional design: why we love (or hate) everyday things*. Basic books, 2004.

[8] J. Allanson and S. H. Fairclough, "A research agenda for physiological computing," *Interacting with Computers*, vol. 16, no. 5, pp. 857–878, 2004.

[9] C. Berka, D. J. Levendowski, M. N. Lumicao, A. Yau, G. Davis, V. T. Zivkovic, R. E. Olmstead, P. D. Tremoulet, and P. L. Craven, "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks," *Aviation, Space, and Environmental Medicine*, vol. 78, no. 1, pp. B231–B244, 2007.

[10] S. H. Fairclough and L. Venables, "Prediction of subjective states from psychophysiology: A multivariate approach," *Biological Psychology*, vol. 71, no. 1, pp. 100–110, 2006.

[11] A. T. Pope, E. H. Bogart, and D. S. Bartolome, "Biocybernetic system evaluates indices of operator engagement in automated task," *Biological Psychology*, vol. 40, no. 1, pp. 187–195, 1995.

[12] P. Hancock and J. Szalma, "The future of neuroergonomics," *Theoretical Issues in Ergonomics Science*, vol. 4, no. 1-2, pp. 238–249, 2003.

[13] F. G. Freeman, P. J. Mikulka, L. J. Prinzel, and M. W. Scerbo, "Evaluation of an adaptive automation system using three EEG indices with a visual tracking task," *Biological Psychology*, vol. 50, no. 1, pp. 61–76, 1999.

[14] M. E. Smith and A. Gevins, "Neurophysiologic monitoring of mental workload and fatigue during operation of a flight simulator," in *Defense and Security*. International Society for Optics and Photonics, 2005, pp. 116–126.

[15] S. H. Fairclough, L. Venables, and A. Tattersall, "The influence of task demand and learning on the psychophysiological response," *International Journal of Psychophysiology*, vol. 56, no. 2, pp. 171–184, 2005.

[16] A. Holm, K. Lukander, J. Korpela, M. Sallinen, and K. M. Müller, "Estimating brain load from the EEG," *The Scientific World Journal*, vol. 9, pp. 639–651, 2009.

[17] F. Yamada, "Frontal midline theta rhythm and eyeblinking activity during a vdt task and a video game: Useful tools for psychophysiology in ergonomics," *Ergonomics*, vol. 41, no. 5, pp. 678–688, 1998.

[18] A. Kamzanova, G. Matthews, A. Kustubayeva, and S. Jakupov, "EEG indices to time-on-task effects and to a workload manipulation (cueing)," *World Academy of Science, Engineering and Technology*, vol. 80, pp. 19–22, 2011.

[19] M.-K. Kim, M. Kim, E. Oh, and S.-P. Kim, "A review on the computational methods for emotional state estimation from the human EEG," *Computational and Mathematical Methods in Medicine*, vol. 2013, 2013.

[20] V. Miskovic and L. A. Schmidt, "Cross-regional cortical synchronization during affective image viewing," *Brain Research*, vol. 1362, pp. 102–111, 2010.

[21] L. Aftanas, A. Varlamov, S. Pavlov, V. Makhnev, and N. Reva, "Affective picture processing: event-related synchronization within individually defined human theta band is modulated by valence dimension." *Neuroscience Letters*, vol. 303, no. 2, pp. 115–118, 2001.

[22] M. Wyczesany, S. J. Grzybowski, R. J. Barry, J. Kaiser, A. M. Coenen, and A. Potoczek, "Covariation of EEG synchronization and emotional state as modified by anxiolytics," *Journal of Clinical Neurophysiology*, vol. 28, no. 3, pp. 289–296, 2011.

[23] D. Nie, X.-W. Wang, L.-C. Shi, and B.-L. Lu, "EEG-based emotion recognition during watching movies," in *5th International IEEE/EMBS Conference on Neural Engineering*. IEEE, 2011, pp. 667–670.

[24] M. Balconi and G. Mazza, "Brain oscillations and bis/bas (behavioral inhibition/activation system) effects on processing masked emotional cues.: Ers/erd and coherence measures of alpha band," *International Journal of Psychophysiology*, vol. 74, no. 2, pp. 158–165, 2009.

[25] M. Balconi and C. Lucchiari, "Eeg correlates (event-related desynchronization) of emotional face elaboration: A temporal analysis," *Neuroscience Letters*, vol. 392, no. 1, pp. 118–123, 2006.

[26] R. J. Davidson, "Anterior cerebral asymmetry and the nature of emotion," *Brain and Cognition*, vol. 20, no. 1, pp. 125–151, 1992.

[27] I. H. Gotlib, "Eeg alpha asymmetry, depression, and cognitive functioning," *Cognition & Emotion*, vol. 12, no. 3, pp. 449–478, 1998.

[28] M. M. Müller, A. Keil, T. Gruber, and T. Elbert, "Processing of affective pictures modulates right-hemispheric gamma band EEG activity," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1913–1920, 1999.

[29] N. Martini, D. Menicucci, L. Sebastiani, R. Bedini, A. Pingitore, N. Vanello, M. Milanesi, L. Landini, and A. Gemignani, "The dynamics of EEG gamma responses to unpleasant visual stimuli: From local activity to functional connectivity," *NeuroImage*, vol. 60, no. 2, pp. 922–932, 2012.

[30] N. Jatupaiboon, S. Pan-ngum, and P. Israsena, "Emotion classification using minimal EEG channels and frequency bands," in *10th International Joint Conference on Computer Science and Software Engineering*. IEEE, 2013, pp. 21–24.

[31] T. D. Pham and D. Tran, "Emotion recognition using the emotiv epoc device," in *Neural Information Processing*. Springer, 2012, pp. 394–399.

[32] L. E. Nacke, "Wiimote vs. controller: Electroencephalographic measurement of affective gameplay interaction," in *Proceedings of the International Academic Conference on the Future of Game Design and Technology*. ACM, 2010, pp. 159–166.

[33] N. Jatupaiboon, S. Pan-ngum, and P. Israsena, "Real-time EEG-based happiness detection system," *The Scientific World Journal*, vol. 2013, 2013.

[34] M. Duvinage, T. Castermans, T. Dutoit, M. Petieau, T. Hoellinger, C. D. Saedeleer, K. Seetharaman, and G. Cheron, "A p300-based quantitative comparison between the emotiv epoc headset and a medical EEG device," *Biomedical Engineering*, vol. 765, pp. 2012–764, 2012.

[35] E. T. Esfahani and V. Sundararajan, "Classification of primitive shapes using brain–computer interfaces," *Computer-Aided Design*, vol. 44, no. 10, pp. 1011–1019, 2012.

[36] S. Fok, R. Schwartz, M. Wronkiewicz, C. Holmes, J. Zhang, T. Somers, D. Bundy, and E. Leuthardt, "An EEG-based brain computer interface for rehabilitation and restoration of hand control following stroke using ipsilateral cortical physiology," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE, 2011, pp. 6277–6280.

[37] C. Vi and S. Subramanian, "Detecting error-related negativity for interaction design," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 493–502.

[38] E. W. Anderson, K. C. Potter, L. E. Matzen, J. F. Shepherd, G. A. Preston, and C. T. Silva, "A user study of visualization effectiveness using EEG and cognitive load," in *Computer Graphics Forum*, vol. 30, no. 3. Wiley Online Library, 2011, pp. 791–800.

[39] T. McMahan, I. Parberry, and T. Parsons, "Modality specific assessment of video game player's experience using the emotiv," *Entertainment Computing*, vol. 7, pp. 1–6, 2015.

[40] O. Sourina and Y. Liu, "A fractal-based algorithm of emotion recognition from EEG using arousal-valence model." in *BIOSIGNALS*, 2011, pp. 209–214.

[41] L. Pessoa, "On the relationship between emotion and cognition," *Nature Reviews Neuroscience*, vol. 9, no. 2, pp. 148–158, 2008.

[42] H. C. Cromwell and J. Panksepp, "Rethinking the cognitive revolution from a neural perspective: How overuse/misuse of the term 'cognition'and the neglect of affective controls in behavioral neuroscience could be delaying progress in understanding the brainmind," *Neuroscience & Biobehavioral Reviews*, vol. 35, no. 9, pp. 2026–2035, 2011.

[43] F. Lotte, M. Congedo, A. Lécuyer, and F. Lamarche, "A review of classification algorithms for EEG-based brain–computer interfaces," *Journal of Neural Engineering*, vol. 4, 2007.

[44] T. Meat, "Super meat boy," 2010.

[45] S. Egenfeldt-Nielsen, J. H. Smith, and S. P. Tosca, *Understanding video games: The essential introduction*. Routledge, 2013.

[46] V. Corporation, "Half life 2," 2004.

[47] Y. Mahajan and G. McArthur, "Does combing the scalp reduce scalp electrode impedances?" *Journal of Neuroscience Methods*, vol. 188, no. 2, pp. 287–289, 2010.

[48] A. Nijholt, D. P.-O. Bos, and B. Reuderink, "Turning shortcomings into challenges: Brain–computer interfaces for games," *Entertainment Computing*, vol. 1, no. 2, pp. 85–94, 2009.

[49] N. Bianchi-Berthouze, W. W. Kim, and D. Patel, "Does body movement engage you more in digital game play? and why?" in *Affective Computing and Intelligent Interaction*. Springer, 2007, pp. 102–113.

[50] S. Salenius, K. Portin, M. Kajola, R. Salmelin, and R. Hari, "Cortical control of human motoneuron firing during isometric contraction," *Journal of Neurophysiology*, vol. 77, no. 6, pp. 3401–3405, 1997.

[51] E. W. Anderson, G. A. Preston, and C. T. Silva, "Towards development of a circuit based treatment for impaired memory: A multidisciplinary approach," in *3rd International IEEE/EMBS Conference on Neural Engineering*. IEEE, 2007, pp. 302–305.

[52] G. Chanel, J. J. Kierkels, M. Soleymani, and T. Pun, "Short-term emotion assessment in a recall paradigm," *International Journal of Human-Computer Studies*, vol. 67, no. 8, pp. 607–627, 2009.

[53] S. Giraldo and R. Ramirez, "Brain-activity-driven real-time music emotive control," in *Proceedings of the 3rd International Conference on Music & Emotion (ICME3), Jyväskylä, Finland, 11th-15th June 2013. Geoff Luck & Olivier Brabant (Eds.). ISBN 978-951-39-5250-1*. University of Jyväskylä, Department of Music, 2013.

[54] A. Bashashati, M. Fatourechi, R. K. Ward, and G. E. Birch, "A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals," *Journal of Neural Engineering*, vol. 4, no. 2, p. R32, 2007.

[55] P. C. Petrantonakis and L. J. Hadjileontiadis, "Emotion recognition from EEG using higher order crossings," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 186–197, 2010.

[56] M. Singh, M. M. Singh, and N. Singhal, "Emotion recognition along valence axis using naïve bayes classifier," *International Journal of Information Technology & Knowledge Management*, vol. 7, no. 1, pp. 51–55, 2013.

[57] Y.-P. Lin, C.-H. Wang, T.-P. Jung, T.-L. Wu, S.-K. Jeng, J.-R. Duann, and J.-H. Chen, "EEG-based emotion recognition in music listening," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 7, pp. 1798–1806, 2010.

[58] M. Murugappan, "Human emotion classification using wavelet transform and KNN," in *2011 International Conference on Pattern Analysis and Intelligent Robotics*, vol. 1. IEEE, 2011, pp. 148–153.

[59] M. Li and B.-L. Lu, "Emotion classification based on gamma-band eeg," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*. IEEE, 2009, pp. 1223–1226.

[60] N. A. Badcock, P. Mousikou, Y. Mahajan, P. de Lissa, J. Thie, and

[60] N. A. Badcock, P. Mousikou, Y. Mahajan, P. de Lissa, J. Thie, and G. McArthur, "Validation of the emotiv epoc® EEG gaming system for measuring research quality auditory erps," *PeerJ*, vol. 1, p. e38, 2013.

[61] L. E. Nacke, S. Stellmach, and C. A. Lindley, "Electroencephalographic assessment of player experience: A pilot study in affective ludology," *Simulation & Gaming*, 2010.

[62] M. Salminen and N. Ravaja, "Oscillatory brain responses evoked by video game events: The case of Super Monkey Ball 2," *CyberPsychology & Behavior*, vol. 10, no. 3, pp. 330–338, 2007.

T. D. Parsons, Ph.D. is Director of the Clinical Neuropsychology and Simulation (CNS) Laboratory and Associate Professor of Psychology at the University of North Texas. He is a leading scientist in neuropsychology and technology. He has been PI of 17 funded projects and an investigator on an additional 13 funded projects. In addition to his patent for eHarmony.com's Matching System (# 6735568), he has invented multiple virtual reality-based neurocognitive assessments. He has over 100 publications in peer-reviewed journals and other fora.

T. McMahan received the M.S. degree in computer science from the University of North Texas at Denton Texas in 2013, where he is currently working towards the Ph.D. degree in computer science and engineering. His research interests are brain computer interfaces, video game testing, gamer experience, video game metrics and artificial intelligence.

I. Parberry, Ph.D. is Director of the Laboratory for Recreational Computing and Professor in the Department of Computer Science & Engineering, University of North Texas. He is author of six books and over 100 articles in peer-reviewed journals, conferences, and other fora on computing topics including algorithms, AI, computational complexity, computer graphics, entertainment computing, game programming, puzzles, neural networks, parallel computing, procedural content development, sorting networks, and video game assessment.