Classification of Video Game 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 manner. Whilst establishing the optimal relation among frequency bands, task engagement, and arousal states is a goal 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, a commercial-grade electroencephalographic (EEG; Emotiv EPOC) headset was used to collect signals from 30 participants. The EEG data was 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 negative (e.g., character death) game-based events. The identification of general gameplay events is best identified using kNN and the Beta band. Results from this study suggest that a combination of classifiers is preferable over selection of a single classifier.**

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

I. Introduction

NEUROGAMING platforms offer potential for supporting the growing body of literature on cognitive assessment of video gamers [1]. Recent developments in neurogaming have involved the use of off-the-shelf brain computer interfaces (BCIs) to interface with well-known games such as “Pacman” [2], “Tetris” [3], and “World of Warcraft” [4]. Furthermore, electroencephalography (EEG) technologies are being utilized increasingly for assessment of user experience while playing with well-known video games. This is accomplished via the recording of brain activity that is then mapped to various game commands. This approach allows for near real-time decoding of a user’s neurocognitive and/or affective (i.e., emotional) state. Interpretation of these states requires identification of an optimal relationship among frequency bands indicating task engagement and arousal. Although this is a common endeavor of neurogaming protocols, a standardized method has yet to be established.

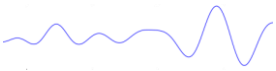
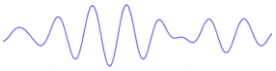
A. Knowledge of User-State During Video Gameplay

The EEG-based BCIs used in neurogaming can provide knowledge of user-state during video gameplay that is important for development and assessment of video game design [5]. Each user has dissimilar reactions to the game they are playing, and without a measure that can be administered on-line, researchers will have limited data for detecting the sources of these variances. The EEG data from BCIs provide signals that are continuously available [6] 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 [7], stress levels [8], and task engagement [9], among others. EEG-based arousal and engagement indices can be gleaned from various sensors continuously, which further increases experimental control [10].

B. EEG for Establishing Indices of Engagement

Assessment of arousal and affective states have been assessed (in the frequency domain) using spectral power in various frequency bands [11]. Frequently analyzed bands include: Delta (0-4 hertz); Theta (4-7 hertz); Alpha 8-12 hertz; Beta (12-30 hertz); and Gamma (>30 hertz). Increased coherence in brain electrical activity (Beta bandwidth) was found when participants viewed affective stimuli that were highly arousing [12]. Delta is seen normally in adults in slow wave sleep and tends to be the highest in amplitude and the slowest waves. Theta can be seen in relaxed and meditative states. Theta power event-related synchronization studies have been found to modulate during transitions in affective state [13]. Interactions between pairs of EEG oscillations (e.g., phase synchronization; coherence) have also been found to be related to affective arousal [14]. Alpha can be found when persons relax and tends to attenuate with cognitive exertion. It has been suggested that higher frequency bands may have greater contribution contribute to arousal response than lower frequency bands [15]. Moreover, Alpha power is related to negative and positive valence states [16] and with distinct affective states [17]. Alpha power frontal asymmetry has been repeatedly reported as a steady correlate of valence [18]. Furthermore, Alpha asymmetry may be related to approach/avoidant affects [19]. Beta of low amplitude is associated with cognitive processes such as thinking and concentration. Gamma rhythms are often considered to represent the binding of various neural networks that perform specific cognitive functions (see Table I). Gamma has been found to reflect changes in affect—event-related synchronization and desynchronization have been associated with affective states [20]. Furthermore, increased gamma phase synchronization has been found to be associated with viewing aversive visual stimuli [21].

TABLE I. Common EEG frequency bands

Frequency Band	Frequency Bandwidth	Filtered Bandwidth	Associated State
Delta	0-4 Hz		Sleep
Theta	4-7 Hz		Drowsy
Alpha	8-12 Hz		Relaxed
Beta	12-30 Hz		Engaged
Gamma	> 30 Hz		Stressed

The use of EEG for assessment of task engagement has been performed in a number of studies using medical grade EEG devices [9], [22]-[25]. For example, Pope and colleagues [9] developed a system to control task automation level relative to whether the user experienced increasing or decreasing engagement. In a further development, Freeman and colleagues [22] extended the measurement of task engagement to include evaluation of performance on each task with absolute values of engagement. Furthermore, increased Theta and decreased Alpha have been found to be correlated with increased number of tasks along with the amount of time a person is awake [23], [24]. Kamzanova and colleagues [25] used a resource-limited vigilance task to examine the sensitivity of various EEG indices. Their findings suggested that increased Theta and lower frequency Alpha can be used for identifying decreases in task engagement. Electroencephalography has also been used for assessment of engagement while users experienced interactive simulations. For example, when participants completed demanding tasks in a flight simulator frontal Theta response increased and parietal Alpha decreased [26]. Higher Theta activity has also been found during demanding tasks occurring while children played video games [27].

C. Neurogaming Using Emotiv EPOC EEG

The Emotiv EPOC 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 [28]–[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 record a gamer's brainwave signature. The system has been found to work well for detection of focused thoughts [32], [33], creativity [34], feelings about avatars [35], and for developing affective ontologies [36]. Comparison of the Emotiv EEG system to the Advance Neuro Technology (ANT) acquisition system during a P300 speller assessment revealed that while the Emotiv EEG was less accurate than expensive ANT systems (medical grade devices), it was able to capture EEG signals appropriate for neurogaming

Researchers have also investigated various EEG processing algorithms to assess Emotiv EEG classifications in controlled laboratory experiments. These have included Emotiv EEG classification during a controlled task in which shapes were being visualized [38]; a controlled task in which hand movement intentions were being detected on the same side of the brain as the hand [39], a laboratory controlled picture presentation task in which positive and negative emotions were elicited [40]; and laboratory based tasks developed for evaluation of cognitive workload [41]. In addition to task engagement, affective states have been measured using Emotiv EEG while users watched a film [42]. The Emotive has the advantage of being easy to setup and noninvasive to the user. As such, the Emotiv EEG offers an instrument that is user-friendly for game developers. While these controlled laboratory tasks have been helpful for validating the level of efficacy found in Emotiv EEG indices, they add less to the understanding of dynamic game events. McMahan et al. [43] were able to find significant difference in the Beta and Gamma bands among various stimulus modalities. Furthermore, increased power estimates were found during high intensity game play (e.g., character death during gameplay). These findings indicate the potential of Emotiv EEG recordings for assessment of gamer experience.

Findings from controlled laboratory tasks regarding the relations between affective and cognitive correlates of brain processing are uncovering the strong implication of cognitive processes in emotions [44]. This has resulted in increasing emphases upon affective neuroscience [45] and the potential for EEG data to proffer valuable information about the participants' 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 the interplay of 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 [46].

Event-related potentials (ERPs) are brain-generated electrical potentials that neurogaming researchers can use to associate neural firings with specific gaming events. Using the Emotiv, ERPs can be recorded noninvasively and ERPs can proffer data specific to a broad range of cognitive and affective processes [37], [65]. Detection of ERPs from specific game events is important for the operation of many stimulus-driven neurogaming systems. While the low strength of the ERP signal compared to the noise (due to artifacts) makes this a challenging

signal detection problem, researchers have shown that distinction between stimulus-response pairings can be made on a single ERP basis [47], [48].

As neurogaming systems increase in use, new properties will need to be taken into consideration. A common difficulty encountered in neurogaming research is the dearth of published objective comparisons of gaming events among classifiers. While there is a great deal of work that has been done looking at EEG and laboratory controlled tasks, little has done to find the best predictors for gaming events. For researchers interested in extending assessment findings to the development of brain-computer interfaces with video games, there is need for the development of game-specific classifiers. In this paper, an approach to developing classifiers of game- based task engagement and affective state estimations 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 extend the arousal and engagement indices found in laboratory-controlled tasks to game-based classifiers that can be used to assess various levels of gaming experience using the Emotiv EEG.

II. METHODS

A. Participants

Thirty young adults were recruited from undergraduate and graduate student pools to take part in the study. Demographic breakdowns were as follows: average age = 20.87 (range 18 to 43); 66 percent of the participants were female; and education levels ranged from 13 to 20 years. Ethnicity distribution was as follows: N=20 Caucasians, N=1 African American, N=4 Hispanic, N=1 Native American, and N=4 Asian Pacific participants. All participants were right handed. All participants endorsed average or greater computer skills. The range of game playing skills was from occasional game play to playing every day. Moreover, all participants reported that they had used a computer at least once per day (every day) in the last year.

Computer experience results revealed that 66 percent rated themselves as experienced, 27 percent as somewhat experienced, and 7 percent as very experienced. Participants reported expending time play games using cell phones (M = 3.47), personal computers (M = 3.47), and consoles (M = 2.3). Sample homogeneity was found for age, education, ethnicity, and sex. No participants reported symptoms of depression or sleepiness. No significant variance was found among participants for game play experience or computer use. All participants received class credit for their participation in the study.

B. Apparatus and Measures

1) Laboratory-Controlled Tasks:

a) *Two-Picture Cognitive Discrimination Task:* In order to assess participants task engagement in the absence of arousal, a laboratory-controlled "Two-Picture Cognitive Task" was used. Participants were instructed to identifying any differences between a pair of color pictures (depicting a landscape). The participants were not informed that the pictures were identical. During this task EEG data was gathered to establish a cognitive marker of attentional processing.

b) *Spider Jump Arousal Stimulus:* To assess user arousal, a laboratory-controlled "Spider Jump Arousal" stimulus was used. The "Spider Jump Arousal" stimulus was first story-boarded and designed on paper. A 3-D model of a venomous head crab was taken from the Half-Life 2 game [49]. The crab was stimulus was chosen for its negatively valanced behavior. The head crab rapidly leaps toward the participant while releasing an angry squeal. The participants were subjected to the Spider Jump Arousal stimuli for 3 seconds without any cue or knowledge that it would occur. During this exposure EEG data was gathered to establish a startle marker of arousal processing.

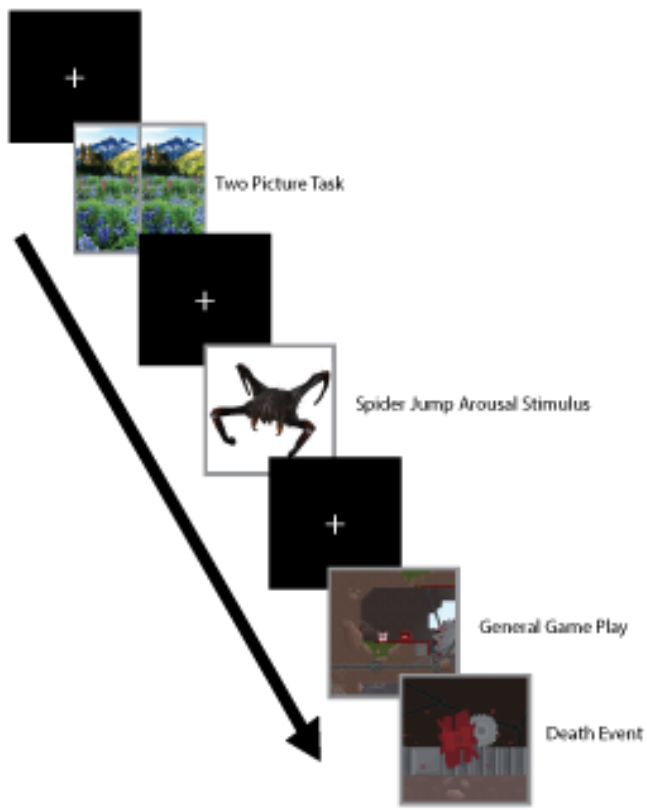


Fig. 1. Chronological order of events that participants encountered.

2) Super Meat Boy:

Participants played a game called *Super* Meat Boy [50], [51] in which they control a character shaped like a cube of meat that jumps around the game level in an effort to avoid saw blades while trying to rescue "bandage girl." A

minimum number of keys (arrow keys and space bar) are used to control the character. As a result, any level of gamer should have the capacity for achieving success. The two primary game events are successful completion of a level and character death. The death of a character occurs when Meat Boy runs into spinning saw blades or falls into fire. Progression through game levels presents the participant with increasing levels of difficulty (each level adds increased numbers of saw blades and larger jumps). A metric of the game is duration needed to get through each level. For this study, the primary game events of interest included: 1) Death events; and 2) "General Game Play". As described above, a "Death event" occurs when the participant's character died. "General Game Play" was sampled during periods in which the player had not experienced any death events for one minute. We followed McMahan and colleagues' [43] use of "general gameplay" events and "high intensity game play" (e.g., character death during gameplay) events because they have been shown to result in significant differences in the Beta and Gamma bands among various stimulus modalities. Furthermore, increased power estimates were found during high intensity game play (e.g., character death during gameplay). These findings indicate the potential of Emotiv EEG recordings for assessment of gamer experience.

3) Emotiv EPOC EEG: The EEG system used in this study was the neurogaming headset found in the Emotiv EEG system: 14 electrodes (saline sensors) that are positioned at AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1, O2; and two sensors (CMS/DRL) for the left and right hemisphere of the head). The letters "F", "P", "T", and "O" above refer to specific lobes of the brain. The brain lobes are as follows: frontal lobe (i.e., F); parietal lobe (i.e., P); temporal lobe (i.e., T): and occipital lobe (i.e., O). F3 and F4 (including AF3 and AF4) are typically associated with attentional processing. FC5, and FC6 are associated with motor processes. F7 is associated with verbal expression and F8 is associated with affective processing. T7 and T8 are associated with memory. P7 and P8 are associated with the participants spatial map. O1 and O2 are associated with visual processing. The 14 data channels organized spatially according to the International 10–20 systems (see Fig. 1). Sampling rate is 128Hz. The bandwidth is 0.2- 45Hz. The digital notch filters are at 50Hz and 60Hz.

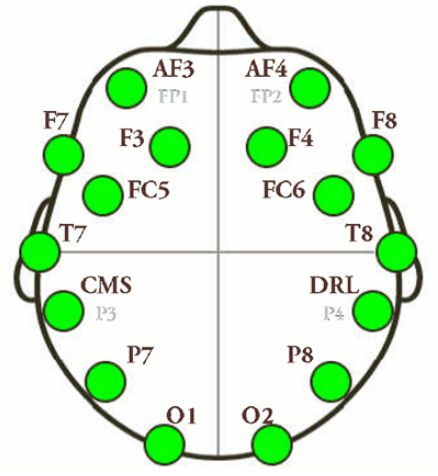


Fig. 2. Sensor locations on Emotiv headset.

C. Procedure

At the start of the testing session, each participant was given an informed consent to read and sign. The informed consent included a waiver to record the participant during the study. The experimenter combed the participants on the right, mid-line, and left sides of their scalp in a firm manner to reduce electrode impedances [52]. After cleaning the pertinent areas on the face and mastoids, the Emotiv EEG headset was positioned on the participant's head. Impedances were verified by the examiner. Each participant played the game while seated in a comfortable chair that was always placed in the same room location. The game was displayed on a Samsung 60-inch plasma screen monitor. Room overhead lights were turned off to reduce glare.

A baseline for each participant was established using a video. Participants were told "to relax and try not to think about anything" as they watched the video. For the first 2:00 minutes the video was blank to establish each participants' minimum brain wave activity. Next, cognitive task engagement was assessed as the video presented the participant with two pictures depicting a landscape (i.e., Two Picture Cognitive Task) and instructed to scan both pictures to identify any differences. This was followed by an assessment of startle response using the spider jump stimuli described above. Afterwards, the screen returned to blank for 1:30 seconds to allow the participant to return to a steady state. In addition to the video exposure, each participant played Super Meat Boy (see Fig. 3). During the first few levels participants were given aid by the researcher to allow the player to become acquainted with the rules and game controls. Participants played Super Meat Boy for 15 minutes. Task presentation order was counter-balanced across participants.

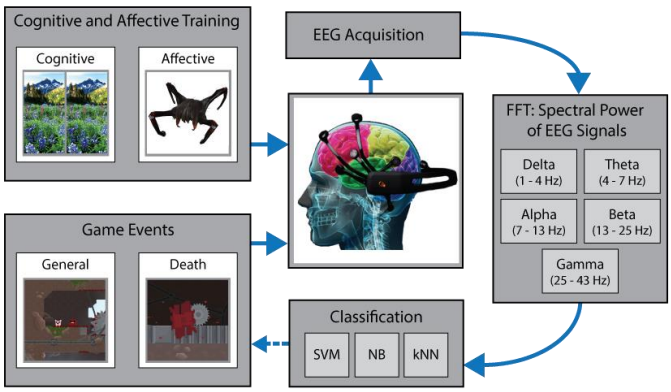


Fig. 3. Flow chart depicting the procedure used to obtain results. The dashed line between classification and game events signifies the eventual use of a closed loop system.

A Hauppauge video capturing device was used to capture each participant's game play (1080p HD; 60 frames per second). This allowed the game play to be synced with the EEG data. A Logitech 9000 HD webcam was used to aid the isolation of events (facial or body movements) that could have impacted the EEG data. The same computer was used to record the EEG and video data. All non-essential programs were closed.

TABLE II. Database Design

ID	EEG Two-Picture Task	EEG Spider Jump	EEG General Game Play	EEG Death Event	F3	F4	Calculated Engagement	Calculated Arousal
1	α β Θ Γ	α β Θ Γ	α β Θ Γ	α β Θ Γ	α β	α β	β/(α+Θ)	(βF3+ βF3) / (αF3+ α F3)
…	FFT Data	FFT Data	FFT Data	FFT Data	FFT Data	FFT Data	…	…
30	FFT Data	FFT Data	FFT Data	FFT Data	FFT Data	FFT Data	β/(α+Θ)	(βF3+ βF3) / (αF3+ α F3)

OpenViBE drift correction, was used to achieve a 128 Hz sample rate. This minimized any syncing issues between the EEG and the video recording of game play. Synchronization of EEG recording software with video recordings involved the use of screen captures before and after every section of the study (baseline video and game play). A time stamp was produced for each screen shot to establish the location of the start and end of each section. Each screen shots was saved for later reference at the data analysis phase.

D.Data Analytics

SAS version 9.1 was used for all analyses. Descriptive statistics were calculated for EEG results and participant demographics. Multiple imputation was used for missing data. OpenViBE and the Emotiv TestBench were used to capture the raw EEG output. Next, EEG data were segmented into epochs that started 100 ms before stimulus onset (0 ms), and ended 750 ms after stimulus onset. Epochs were calculated for 4 different modalities: 1) two-picture cognitive task; 2) spider jump arousal stimulus; 3) death events; and 4) general game play (see Table II).

Artifacts (e.g., blinking, head movements, body movements) can cause unwanted data in EEG data. While most EEG analyses require artifact removal to help identify medical issues, these types of artifacts are common in everyday game play [53] and they can be used for further analysis as body movement or other movement can signify engagement [54]. That said the EEG artifact data were annotated as artifacts in situations where noticeable deflection in the EEG were observed. Eye blinks and other muscle movements (in addition to physical movement of the sensors themselves) artifacts were removed before the EEG traces were processed. The Emotiv SDK was used to automatically detect and records eye blinks. Since muscle contraction and control are mostly governed outside of the frequency range of interest [55], we used frequency band limiting procedures (e.g., low-pass, high-pass and notch filters) to remove these signal components. After removing EEG artifacts, the energy densities of the alpha or theta frequency bands were assessed to see if there were any trials with changes greater than 20 percent of their original values [56]. No trials in this study had excessive signal degradation from movement or excessive change in spectral densities.

The EEG spectral power in various bands has been found to be correlated with emotions [57]. Power estimates (µV2) were established using a fast Fourier transform (FFT) and a one second hamming window without overlap for Delta (1 – 4 Hz), Theta (4- 7 Hz), Alpha (7 -13 Hz), Beta (13– 25 Hz) and Gamma (25 – 43 Hz). This was done for all 14-sensor location on the Emotiv EEG headset. While some EEG studies have channels ranging from 32 channels (for routine exams) up to 256 channels (for source localization) and sample at up to 1000Hz, the Emotiv EEG 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. The baseline and stimulus signals were transformed to determine the power change and frequency shift induced by the task [41]. These values were used to calculate the cognitive load from the 14 sensors. Spatial averaging of the 14 values resulted in a single measurement for analyses. The data were normalized with the natural logarithm (ln).

Prior research has developed an engagement index can be calculated from the ratio of Beta / (Alpha + Theta) EEG bands [9], [22]. Moreover, the engagement index has been found to be associated with information-gathering, visual scanning and sustained attention [7]. For this study, an engagement index was calculated for each participant using an aggregate measurement form all sensors. In addition to engagement, an arousal index has been developed via the following (BetaF3 + BetaF4) / (AlphaF3 + AlphaF4) [58].

All user data sets were analyzed together to build 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 14 EEG measurements each for Alpha, Beta, Theta, Gamma, along with the calculated signals global Alpha (see Fig. 4), global Beta (see Fig. 5), global Theta (see Fig. 6), Engagement Index (see Fig. 7), and Arousal Index (see Fig. 8). Using the data from each participant 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.

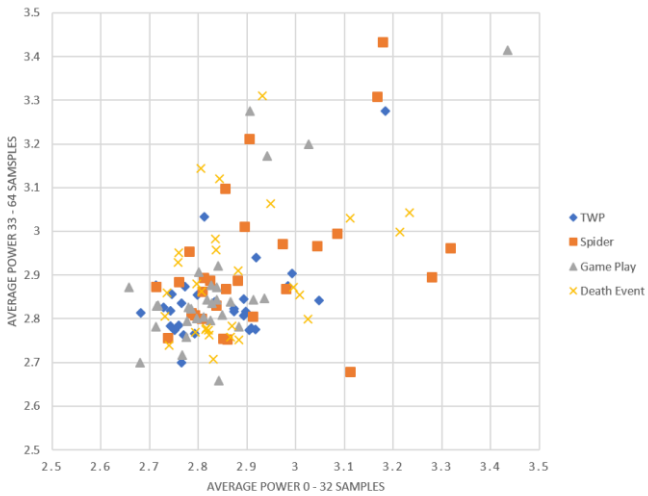


Fig. 4. Scatter plot depicting data points from the Alpha signal for Two-Picture Cognitive task, Spider Jump Arousal Stimulus, General Game Play, and Death Events.

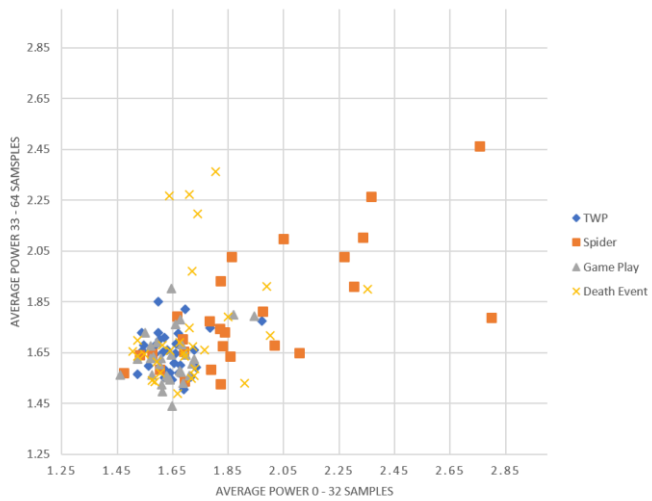


Fig. 5. Scatter plot depicting data points from the Beta signal for Two- Picture Cognitive task, Spider Jump Arousal Stimulus, General Game Play, and Death Events.

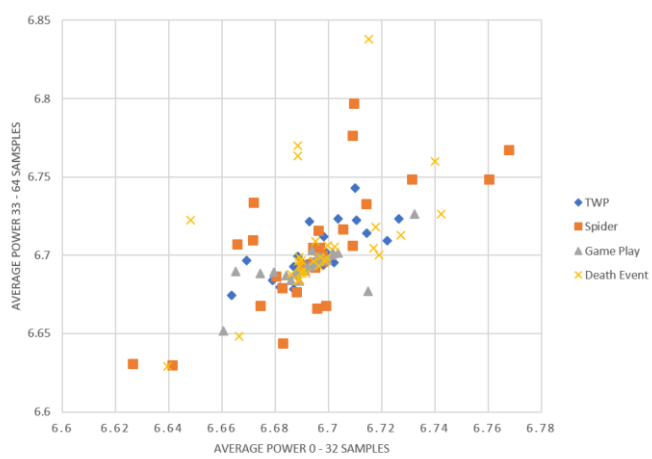


Fig. 6. Scatter plot depicting data points from the Theta signal for Two-Picture Cognitive task, Spider Jump Arousal Stimulus, General Game Play, and Death Events.

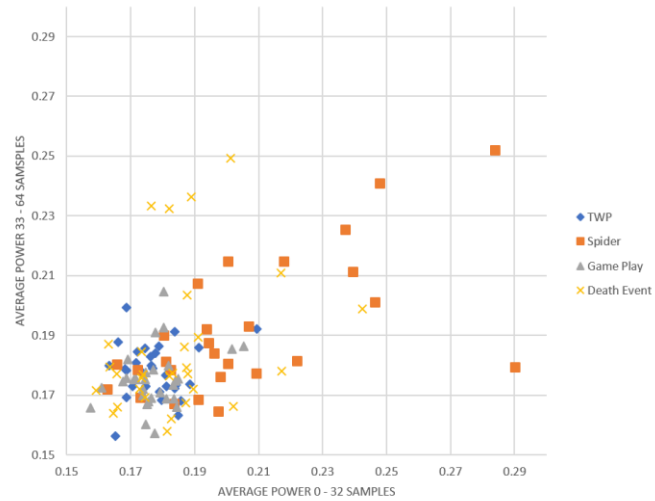


Fig. 7. Scatter plot depicting data points from the Engagement signal for Two-Picture Cognitive task, Spider Jump Arousal Stimulus, General Game Play, and Death Events.

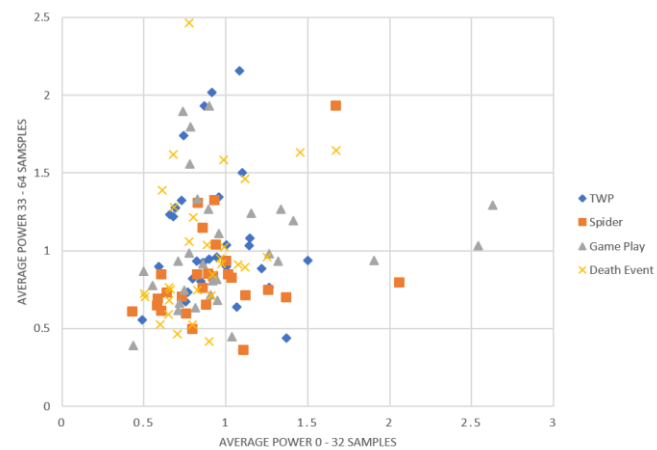


Fig. 8. Scatter plot depicting data points from the Arousal signal for Two-Picture Cognitive task, Spider Jump Arousal Stimulus, General Game Play, and Death Events.

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. The SVM technique has been used for arousal state estimation and results revealed a recognition accuracy of 83 percent [60]. We expected that the SVM would be a good index of arousal related to game play events in general, and for death events specifically. During the training process, 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, 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

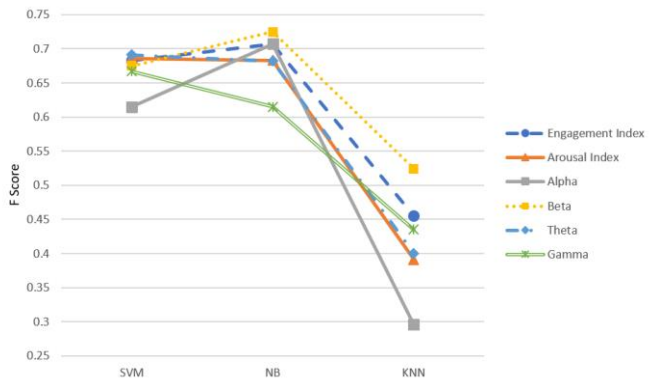


Fig. 9. Overall classifier results for each machine learning algorithm.

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 [59]. 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. When event related potentials are included as a feature, the NB has been used to classify emotions in two classes (low arousal and high arousal) with a classifying accuracy of 56% [61]. In a related analysis, NB has been found to provide recognition accuracy of 70 percent for two classes (as reviewed in Nie et al. [15]). We expected that the NB would be a good index of arousal related to game play events in general, and for death events specifically. This classifier computes the probability that some data points belong to a specific class using (1). To perform the classification, the algorithm chooses the class with the highest probability as its result. 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 maximum likelihood estimates to 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.

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}\right) \quad (1)$$

3) k-Nearest Neighbor: k-Nearest Neighbor (kNN) is a supervised learning algorithm that classifies data into different groups based upon how close it is located to a category. During training the classifier stores each category datum into a feature vector. New data is then classified based upon the training sample that has the shortest distance to the new data point, which is calculated using (2) An issue that can arise from the kNN classifier is if the data does not have an even distribution

$$D(a,b) = \sqrt{\sum_{i=1}^{n}(b_i - a_i)^2} \quad (2)$$

causing the classifier to favor one category over the other. In a study of arousal state estimation Lin et al. [62] 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 [63]. We expected that the kNN would be a good index of arousal related to game play events in general, and for death events specifically.

TABLE III. Overall Classifier Results

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 IV. Signal Classification Results

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 III). 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

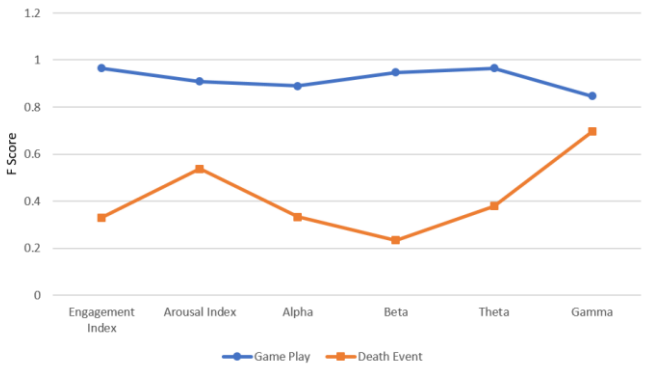


Fig. 10. SVM classifications F-scores across each signal for game play and death events.

TABLE V. Classifiers F Scores

Signal	SVM	NB	kNN
Engagement	0.683	0.707	0.455
Arousal	0.686	0.683	0.391
Alpha	0.615	0.707	0.296
Beta	0.675	0.725	0.524
Theta	0.691	0.682	0.400
Gamma	0.667	0.615	0.435

points to test each classifier (30 for the General Gameplay Events and 30 for the Death Events). The Engagement Index (Beta / (Alpha + Theta); Pope et al. [9] and Freeman et al. [22]), Arousal Index (BetaF3 + BetaF4) / (AlphaF3 + AlphaF4), as well as Alpha, Beta, Theta, and Gamma bands were used as the predictors for each machine learning algorithm. Each predictor was tested individually to identify the strongest signals for classification (see Table IV and Table V).

1) Machine Learning Classifiers and EEG Power Spectral Bands: The strongest classifier was NB especially when using the Theta and Beta signals (Fig. 9 and Table IV show the overall accuracy for each classifier using the different 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: 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 (see Fig. 10). The strongest signals again were Beta, Theta, and the Engagement Index. The Gamma

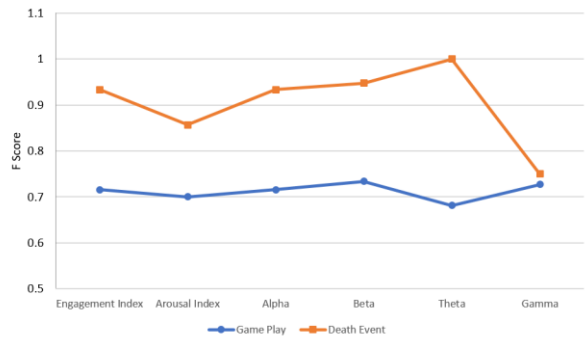


Fig. 11. NB classification F-scores across each signal for game play and death events.

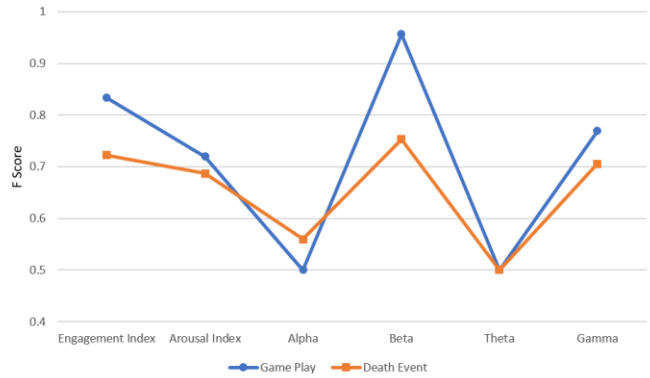


Fig. 12. kNN classification F-Scores across each signal for game play and death events.

band showed the most potential with this classifier as it did the best job in classifying Death Events.

The NB classifier did the best job classifying Death Events from the training data (see Fig. 11). The strongest signals were Theta and Beta followed by Alpha and the 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. 12). 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 than 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 [46]. Specifically, the EEG signals from users were logged as participants experienced various stimulus modalities (both laboratory-controlled tasks and gaming events) 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 [12]. Further, McMahan et al. [43] found significant difference in the Beta band among various stimulus modalities. The Naive Bayes classifier had an overall average of 70 percent 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 [61]. Although the kNN classifier produced the highest accuracy rate with the Beta 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

The laboratory-controlled Two Picture Cognitive Discrimination Task and the Spider Jump Arousal Stimulus to train the SVM classifier, which 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 percent [60]. Again, the strongest signals were Beta, Theta, and the Engagement index. The Gamma band showed the most potential with the SVM classifier, which did the best job in classifying Death Events. Although the Gamma band has been shown in previous research to identify changes in emotion [64], 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 Events from the training data. The strongest signals were Theta and Beta followed by Alpha and the Engagement index. Unlike the SVM classifier, the Gamma band performed the worst. The overall trend of the NB classifier was the most consistent and reliable in distinguishing between General Gameplay Events and Death Events. Although the kNN classifier had greater variability relative to which signal was being used for classification, it did an overall better job with General Gameplay Events. The Beta signal was the strongest predictor for the kNN classifier for both General Gameplay Events and Death Events. Although the Alpha signal was the weakest predictor, it did perform better in predicting Death Events than General Gameplay Events. The overall trend of the kNN classifier was erratic, but potential was observed when using the Beta signal.

IV. Real-time implementation

Findings from the comparison of machine learning classifiers were used to develop a program (i.e., BrainWave) that can be used by researchers and game developers to analyze player experience. BrainWave takes raw EEG signals from the Emotiv EPOC headset as inputs and outputs text to a window. BrainWave is designed to satisfy the following criteria:

1) It must run in real time. Since games and game development tend to move at a fast pace, being able to evaluate a player in real time while they are playing can assist the developer with implementing changes to the game more quickly and efficiently.
2) It must be able to collect data with minimal interference to the player's experience. Utilizing psychophysiological data allows the player to be uninterrupted while they play as well as provide immediate feedback to the researchers.
3) Its output must be user-friendly, that is, presented in terms familiar to the average game developer who may not be an expert in the use of EEG. Most raw psychophysiological data consist of voltage or impedance fluctuations, which usually requires an expert to interpret and understand.

A. Design

Figure 13 illustrates the process of using BrainWave which is broken up into five steps: data capture, data processing, training, testing, and analysis. Each of these is addressed below.

1) Data Capture: Data capture begins with a successful connection to the Emotiv EPOC headset. Utilizing Emotiv's dynamic link library (DLL), a function is called that creates and returns a handle to the headset. Once the handle is created and the connection state verified, data capture begins. As the Emotiv EPOC headset measures raw data, it stores it into a buffer. The size of the buffer determines the amount of data that is collected before it is accessible. The buffer size is determined by the number of epochs the user chooses to collect at one time. BrainWave analyzes the EEG data once per second, thus a buffer is needed that holds exactly one

second worth of EEG data. Once the buffer is full, the Emotiv returns the buffer to the user in the form of an array of structs. The array is made up of all the data from each sensor that is collected during the epoch. Because the headset collects 128 samples a second and has 14 sensors it creates an array that is 128 x 14. This array provides an efficient way to access the collected data so that processing can begin.

2) Data Processing: For each full buffer, the following four steps are performed.

1) Remove potential noise that may impact the data. EEG works by measuring electrical activity at the scalp. However, in the process of measuring the electrical activity direct current (DC) noise may be introduced. This noise comes from the battery that is powering the headset and the radio transmitter. Removing this noise requires the application of a high pass filter to the data that removes any frequency less than 1 Hz and a low pass filter that removes any frequency over 50 Hz. The noise must be removed from each sensor location.
2) Calculate the average voltage across all of the sensors. This is accomplished by adding all of the values from each sensor and dividing it by 14 (i.e., the number of sensors). This calculation is performed for each individual measurement that the headset makes. Averaging the sensor data allows for a global overview of the changes the user is experiencing. It also helps in eliminating any potential outlier data caused by sensor malfunction.
3) Apply the Fast Fourier Transform (FFT) to the data. This step makes the EEG data more meaningful by converting it from time-domain to frequency-domain. The amplitudes for each frequency are calculated by squaring each individual frequency. This is also known as calculating the power estimates. The average is taken of all the frequencies in each band. The resulting value gives the overall power level of the band for the current epoch.
4) Use the band data to calculate the engagement and arousal indices. This is done by plugging the data into the appropriate equations that were tested in the previous sections of this paper.

All of the above steps are repeated each time an epoch of data is gathered by the headset.

3) Data Training: Gathering the data to train the machine learning algorithm first requires gathering the data at the appropriate time. Using video to collect user baseline data for training allows the BrainWave program to know which data it needs to save. The BrainWave controls the start of the baseline video. The BrainWave program uses a timer to identify the points at which training data is to be collected by simply monitoring stimulus presentations relative to the number of seconds from the start of the video.

The BrainWave program then begins using the data to train the three machine learning algorithms. The BrainWave program utilizes a Naïve Bayes (NB) classifier that trains under the assumption that the data has a Gaussian distribution. During the training phase, the mean and variance is calculated and stored for all of the training data. The training of the k-nearest neighbor (k-NN) algorithm requires storing the data into vectors representing the various classes. The support vector machine (SVM) plots the training data to find the best possible linear hyperplane that provides the maximized margin between the data.

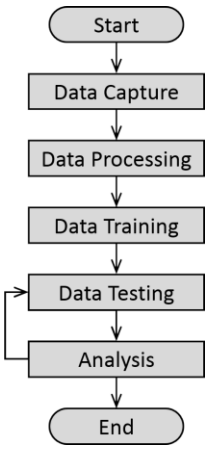


Fig 13 High-level flowchart for the process of measuring player experience.

The engagement and arousal levels are also calculated from the training data. Utilizing the middle base line, the BrainWave program calculates the upper and lower threshold for both engagement and arousal. These thresholds are used during testing to determine if the player has left a state of flow while playing a video game. The concept of "flow" may be understood as an optimal state of engagement that is characterized by a level of optimally focused concentration that results in comprehensive immersion and absorption within an activity [65], [66].

4) Data Testing: After training the classifier, BrainWave switches into testing mode wherein the user is playing the game. During the testing phase, all of the data collected from the headset is processed the same way as described above. As the BrainWave program finishes processing each epoch of data it sends the data to the classifiers to determine if the player is playing the game or has died in the game. For the NB classifier, the BrainWave program uses the mean and variance calculated from the training data to compute the probability that the current epoch of time belongs to a game event or a death event. Depending on which probability is larger, the program selects the appropriate class.

To test using the k-NN, the BrainWave program puts the current epoch of data into a vector. The k-NN classifier then calculates a distance between the test data, game training data, and death training data. The classifier chooses which class to pick based upon the shortest calculated distance. Finally, the SVM tests the data by placing in into the hyperplane established during the training phase. The SVM picks the class closest to the current epoch to which it belongs by identifying the side of the hyperplane on which the data lands.

5) Analysis: The final task for BrainWave is the analysis of results from the classifiers. Each classifier choses which class to which the current epoch belongs. Initially each classifier is independent of each other so that if one classifier predicts a death event has occurred then the BrainWave program outputs a time stamp along with an indication

message. In a different iteration of the program the classifiers are dependent on each

Table VI. Program Classification Results

ML	TP	FP	FN	TN	SEN	SPEC
SVM	25	5	6	564	80.65	99.12
NB	23	20	8	549	74.79	96.49
kNN	25	5	6	564	80.65	99.12
All 3	28	16	4	552	87.50	97.18

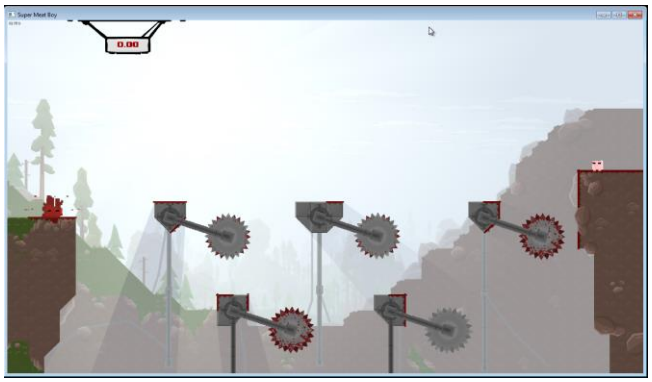


Fig. 14. Screen shot of Super Meat Boy dying and BrainWave showing that it correctly classified a death event using the kNN classifier.

other. The BrainWave program requires that the classifiers must all agree that a death event has occurred before the program will output that a death event has occurred. The program then calculates whether the player has left a state of flow by comparing the current epoch's engagement and arousal level to the upper and lower thresholds. If the player goes above or below the threshold the program timestamps and indicates that this has occurred.

B. User Testing

To validate BrainWave, two participants wore the Emotiv EPOC headset and played Super Meat Boy. Unlike the first study conducted, all of the data processing was completed by BrainWave in real time. Each participant first watched the same video from the previous study to establish their base line and collect the training data. They then were instructed to play Super Meat Boy for 10 minutes.

The first participant went through the first iteration of the program which kept the classifiers independent of each other (see Table VI). This iteration achieved an 83 percent correct classification of death events experienced by the player. This was achieved by both the k-NN and SVM classifiers (see Figure 14 and 15). The NB classifier did not fare as well as it was more sensitive to the player's frustration while playing. One instance of this was when the player said out loud ``I am going to die." The participant did not die but the NB classified this epoch as a death event. An interesting observation was that 75 percent of the time all the classifiers agreed that a death event occurred.

Throughout game play the player moved into and out of flow continuously. From visual and audible observation of the player this constant fluctuation in engagement and arousal could be attributed to frustration.

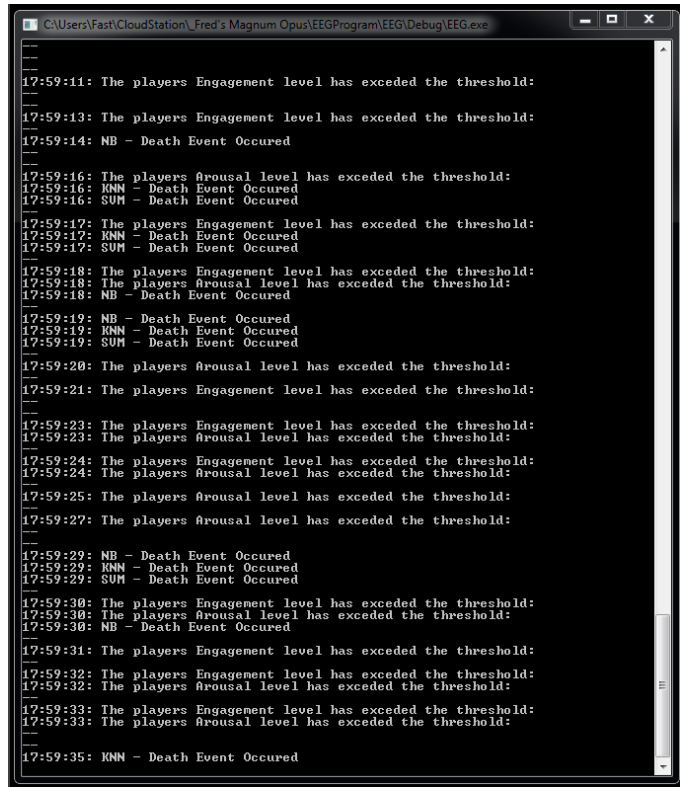


Fig 15. Screen shot of BrainWave showing that it correctly classified a death event using the kNN classifier.

Figure 14 shows several times stamps that represent the constant fluctuation in engagement and arousal in this particular level the player had died 9 times and reported being very upset about dying.

The observation that the classifier agreed 75 percent of the time led to the creation of a second iteration of the program. In this iteration, all of the classifiers had to agree a death event had occurred before the program thought it had occurred. The second participant used the new iteration of the program during their session. The program could obtain a correct classification percentage of 64 percent of the time (see Table VI). This demonstrates that the current best choice for classification is either the k-NN or the SVM algorithms.

The second participant was visually and audibly calmer while playing Super Meat Boy. This was reflected in the second participant's flow during the game. The second participant left the state of flow significantly less than the first participant. The majority of times that the player did leave the state of flow occurred right before or after dying.

C. 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. [67] 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.

It is important to note that we followed McMahan and colleagues' [43] use of "general gameplay" events and "high intensity game play" (e.g., character death during gameplay) events. Future studies may wish to add additional events. For example, researchers may wish to look at game events that are associated with various awards that occur during gameplay. Furthermore, there may be instances where researchers will want to look at game events that facilitate learning and memory.

Hence, 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 [68], [69]. 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.

V.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] 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.

[2] 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.

[3] 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.

[4] 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.

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

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

[7] 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.

[8] 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.

[9] 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.

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

[11] 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.

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

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

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

[19] I. H. Gotlib, "Eeg alpha asymmetry, depression, and cognitive function- ing," Cognition & Emotion, vol. 12, no. 3, pp. 449–478, 1998.

[20] M. M. Mu¨ller, 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.

[21] 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.

[22] 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.

[23] 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.

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

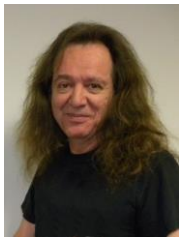
[25] A. Kamzanova, G. Matthews, A. Kustubayeva, and S. Jakupov, "EEG indices to time-on-task effects and to a workload manipulation (cueing)," World Academy of Science, Engineering and Technology, vol. 80, pp. 19–22, 2011.

[26] 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.
[27] F. Yamada, "Frontal midline theta rhythm and eye blinking activity during a vdt task and a video game: Useful tools for psychophysiology in ergonomics," Ergonomics, vol. 41, no. 5, pp. 678–688, 1998.
[28] G. Al-Hudhud, "Affective command-based control system integrating brain signals in commands control systems," Computers in Human Behavior, vol. 30, pp. 535–541, 2014.
[29] 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.
[30] T. D. Pham and D. Tran, "Emotion recognition using the emotiv epoc device," in Neural Information Processing. Springer, 2012, pp. 394–399.
[31] D. Salisbury, S. Driver, and T. Parsons, "Brain–computer interface targeting non-motor functions after spinal cord injury: a case report," Spinal cord, vol. 53, pp. S25–S26, 2015.
[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] K. Muldner and W. Burleson, "Utilizing sensor data to model students' creativity in a digital environment," Computers in Human Behavior, vol. 42, pp. 127–137, 2015.
[35] M. Wrzesien, A. Rodr´ıguez, B. Rey, M. Alcan˜iz, R. M. Ban˜os, and M. D. Vara, "How the physical similarity of avatars can influence the learning of emotion regulation strategies in teenagers," Computers in Human Behavior, vol. 43, pp. 101–111, 2015.
[36] R. Gil, J. Virgili-Goma´, R. Garc´ıa, and C. Mason, "Emotions ontology for collaborative modelling and learning of emotional responses," Computers in Human Behavior, 2015.
[37] 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.
[38] 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.
[39] 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.
[40] C. Vi and S. Subramanian, "Detecting error-related negativity for inter- action design," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2012, pp. 493–502.
[41] 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.
[42] O. Sourina and Y. Liu, "A fractal-based algorithm of emotion recognition from EEG using arousal-valence model." in BIOSIGNALS, 2011, pp. 209–214.
[43] T. McMahan, I. Parberry, and T.D. Parsons, "Modality specific assessment of video game player's experience using the emotiv," Entertainment Computing, vol. 7, pp. 1–6, 2015.
[44] L. Pessoa, "On the relationship between emotion and cognition," Nature Reviews Neuroscience, vol. 9, no. 2, pp. 148–158, 2008.
[45] T.D. Parsons, "Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences," Frontiers in human neuroscience, 660, 9, 1-19, 2015.
[46] F. Lotte, M. Congedo, A. Le´cuyer, and F. Lamarche, "A review of classification algorithms for EEG-based brain–computer interfaces," Journal of Neural Engineering, vol. 4, 2007.
[47] M. K. Ahirwal, A. Kumar, and G. K. Singh, "A New Approach for Utilization of Single ERP to Control Multiple Commands in BCI", International Journal of Electronics Letters, (Taylor & Francis), Vol. 2, No. 2, 2014, pp. 166-171.
[48] M. K. Ahirwal, A. Kumar, and G. K. Singh, "Sub band Classification of Decomposed Single ERP Co-variants for Multi-class BCI: A Qualitative and Quantitative Approach", IET Science, Measurement & Technology, 10, 4, 355-363, 2016.
[49] V. Corporation, "Half-life 2," 2004.
[50] T. Meat, "Super meat boy," 2010.
[51] S. Egenfeldt-Nielsen, J. H. Smith, and S. P. Tosca, Understanding video games: The essential introduction. Routledge, 2013.
[52] 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.
[53] 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.
[54] 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.
[55] 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.
[56] 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.
[57] 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.
[58] S. Giraldo and R. Ramirez, "Brain-activity-driven real-time music emotive control," in Proceedings of the 3rd International Conference on Music & Emotion (ICME3), Jyva¨skyla¨, Finland, 11th-15th June 2013. Geoff Luck & Olivier Brabant (Eds.). ISBN 978-951-39-5250-1. University of Jyva¨skyla¨, Department of Music, 2013.
[59] 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.
[60] 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.
[61] M. Singh, M. M. Singh, and N. Singhal, "Emotion recognition along valence axis using na¨ıve bayes classifier," International Journal of Information Technology & Knowledge Management, vol. 7, no. 1, pp. 51–55, 2013.
[62] 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.
[63] 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.
[64] 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.
[65] M. Csikszentmihalyi, M. "Finding flow. New York: Basic Books. 1997.
[66] T.D., Parsons, T.D., & J. Reinebold. "Adaptive Virtual Environments for Neuropsychological Assessment in Serious Games." *IEEE Transactions on Consumer Electronics, 58*, 197-204, 2011.
[67] 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.
[68] L. E. Nacke, S. Stellmach, and C. A. Lindley, "Electroencephalographic assessment of player experience: A pilot study in affective ludology," *Simulation & Gaming*, 2010.
[69] 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 Computational Neuropsychology and Simulation (CNS) Laboratory and 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 Ph.D. degree in computer science from the University of North Texas at Denton Texas in 2015. He completed a postdoctoral fellowship in the Computational Neuropsychology and Simulation (CNS) lab with Dr. Parsons (PI). He is now on faculty in the Computer Science Department at University of Texas, Dallas. 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.