Evaluating player task engagement and arousal using electroencephalography

Timothy McMahan\textsuperscript{a}, Ian Parberry\textsuperscript{a}, Thomas D. Parsons\textsuperscript{b,*}

\textsuperscript{a}University of North Texas, Department of Computer Science and Engineering, 1155 Union Circle \#311366, Denton, Texas, 76203
\textsuperscript{b}University of North Texas, Department of Psychology, 1155 Union Circle \#311280, Denton, Texas, 76203

Abstract

Advances in affective computing technologies have made it possible for researchers to investigate brain function while users interact in virtual environments. Progress in sensors and algorithms for off-the-shelf EEG systems has made it possible for gaming researchers to perform real-time estimation of human cognitive and affective states using EEG. In this study our aim was to coordinate “Task Engagement” data with “Arousal-Valence” data. The resulting coordinate was designed for application to expressive transformations to video game play in real time by tuning different performance parameters in an Engagement-Arousal rule system. Results revealed that the engagement index (Beta / (Alpha + Theta)) was capable of differentiating high intensity game events (Player Death) from general game play. Given that higher levels of engagement during death events may reflect increase in autonomic response, we also measured arousal by using (BetaF3 + BetaF4) / (AlphaF3 + AlphaF4) and valence using (AlphaF4 / BetaF4) - (AlphaF3 / BetaF3). Results revealed that arousal increases and valence decreases during high intensity game events (Player Death) when compared to lower intensity game events (General Game Play). Given our desire to establish “Task Engagement” data with “Arousal-Valence” coordinates for a flow model, we divided the data into quartiles, which allowed us to establish upper and lower thresholds to indicate when the player has left a state of flow. Our aim was to use an off-the-shelf EEG system to establish “Task Engagement” and “Arousal-Valence” coordinates during video game play that can be used for a flow model. It is believed that this model will allow for future use of the Emotiv for assessing the cognitive and emotional processing of the player.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Peer-review under responsibility of AHFE Conference

Keywords: Electroencephalography; Engagement; Arousal; Video games; Flow; Immersion;

* Corresponding author. Tel.: 940-565-4329.
E-mail address: Thomas.Parsons@unt.edu
1. Introduction

Assessment of player immersion during video gameplay is often a difficult task. Several studies have implemented questionnaires to make assumptions about the player’s experience while in the virtual environments [1-4]. Issues can arise from using these questionnaires including but not limited to wording, context, and format [5]. Given the growing community of video gamers, there is an increasing need for a better approach to indicate when players are immersed at an adequate level to keep them continuously entertained. Neurogaming approaches are emerging that allow for adaptation to fluctuations in engagement [6], cognition [7, 8] and arousal [8-10]. New advances in brain-computer interfaces (BCI) have allowed researchers an inexpensive alternative to laboratory-based systems; these wireless electroencephalographic (EEG) systems offer user metrics for the determination of task engagement and arousal.

1.1. EEG as a metric during video gameplay

Researchers have investigated different EEG processing algorithms to assess classification of positive and negative emotion elicited by pictures [11, 12] and evaluation of cognitive workload [13]. Sourina and Liu [14] used EEG to measure a user’s affective states while they watched film. EEG provides a means of accessing and recording neural activity, allowing a computer to retrieve and analyze information from the brainwave patterns produced by thought. McMahan et al [15] 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. Beta rhythm has been shown to increase with attention and vigilance in general and during video game play specifically [16]. Salmin and Ravajja [16] used EEG to isolate specific game events from the EEG data. Using Super Monkey Ball 2 as their test platform they were able to detect changes in the brain wave bands as different event occurred during game play.

1.2. EEG for establishing indices of engagement and arousal

Using EEG to measure task engagement is not a new concept. Pope et al. [17] built a system to control the level of task automation based upon the whether the operator had increasing or decreasing engagement. Freeman et al. [18] 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. Berka et al. [19] 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. Recently, Kamzanova et al. [20] 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.

In the frequency domain, the spectral power in various frequency bands has been used for assessing arousal and affective states [21]. Beta, EEG coherence has been found to increase when participants viewed highly arousing stimuli [22]. Theta power event-related synchronization studies have found modulation during transitions in affective state [23]. 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 [24]. Researchers emphasize the potential of Alpha power variance with the negative and positive valence states [25] or with discrete affective states such as happiness, sadness, and fear [23]. Gamma power event-related synchronization and desynchronization has been related to affective states such as happiness and sadness [26]. Also, increases in the gamma phase synchronization index have been induced by unpleasant visual stimuli [27].
1.3. EEG to evaluate a state of flow

Flow or immersion can be described as a state in which the player is in a proper level of both engagement and arousal. If the player has too much of either engagement or arousal the player could be in a state of stress were as if it is low the player may have entered a state of boredom. Nacke et.al [28] showed that EEG data could be used to determine player experience across entire level designs. Berta et al [29] used EEG and other physiological signals to assess a player state by using game levels designed to induce boredom, flow, or anxiety. The majority of research has focused on flow across entire levels of play versus measuring changes in immersion throughout a level. The goal of our research is to establish flow model by combining “Task Engagement” data with “Arousal-Valence” data. Using data from the model we established upper and lower thresholds to indicate when the player has left a state of flow.

2. Methods

2.1. Participants

EEG data was collected from 30 healthy participants (66% female, mean age = 20.87, range 18 to 43). Participants were recruited from undergraduate graduate schools; 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). Participants reported they used a computer at least once every day with 30% saying they used the computer several times a day. 66% participants 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, and self-reported symptoms of depression. 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. The participants received class credit for their participation in the study.

2.2. Apparatus

2.2.1. Super Meat Boy

Super Meat Boy [30] 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. 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.

2.2.2. Game Experience Survey

Participants answered a series of questions assessing prior videogame 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.

2.2.3. Emotiv EPOC EEG

This Emotiv has 14 electrodes (saline sensors) locating at AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1, O2 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 conduction. The sampling rate is 128Hz, the bandwidth is 0.2–45Hz, and the digital notch filters are at 50Hz and 60Hz.
2.3. 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 [31]. 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. 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. 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.

2.4. Data Analytics

All data were analyzed using SAS version 9.1. Descriptive statistics were calculated for participant demographics and for EEG results (see Table 1). Missing data were imputed by either mean substitution or last case carried forward. The Emotiv Epoc headset was used to capture the EEG data from each participant. 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 General Game Play and Death events.

<table>
<thead>
<tr>
<th>Indices</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement Index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Game Play</td>
<td>0.33</td>
<td>0.148</td>
<td>0.027</td>
</tr>
<tr>
<td>Death Event</td>
<td>0.45</td>
<td>0.238</td>
<td>0.044</td>
</tr>
<tr>
<td>Arousal Index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Game Play</td>
<td>0.22</td>
<td>0.262</td>
<td>0.005</td>
</tr>
<tr>
<td>Death Event</td>
<td>0.25</td>
<td>0.398</td>
<td>0.007</td>
</tr>
<tr>
<td>Valence Index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Game Play</td>
<td>0.70</td>
<td>0.682</td>
<td>0.124</td>
</tr>
<tr>
<td>Death Event</td>
<td>0.45</td>
<td>0.545</td>
<td>0.099</td>
</tr>
</tbody>
</table>
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 gameplay analysis. These types of artifacts are common in everyday gameplay [32]. These artifacts can actually be used for further analysis as body movement or other movement can signify engagement [33]. The EEG artefact data was annotated as artefact 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 [34], 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 [35] 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 power estimates (μV²) 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 [13] 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.

Pope et al. [17] and Freeman et al. [18] have shown that an engagement index can be calculated by taking the ratio of Beta / (Alpha + Theta) EEG bands. Berka et al. [19] 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) [36].
2.5. Results

We completed a repeated-measures analysis of variance assessment (ANOVA) across the index of engagement, the arousal index and the valence index to verify a difference exists between general game play and death events. Results from the repeated measures ANOVA using the indices as the within subject factor for dependent variable general game play and death events revealed a significant different for the main effect (F(2,28) = 183.22, p < 0.001, partial eta 2 = 0.68). These results represent the difference in the formulas used to calculate each index.

Follow-up test of repeated within-subject contrasts revealed difference in between general game play and death events within each index. The engagement level during death events was significantly increased in comparison to general game play (t(1,29) = 2.720, p = 0.011). The arousal was also significantly increased during death events in comparison to general game play (t(1,29) = 3.959, p < 0.001). The valence index did not yield any significant differences between general game play and death events. However, it did yield an interesting trend that valence usually decreased during death events when compared to general game play events (see table 1). Given our desire to establish “Task Engagement” data with “Arousal-Valence” coordinates for a flow model, we divided the data into quartiles, which allowed us to establish upper (Task Engagement = 0.17, Arousal = 0.24) and lower (Task Engagement = 0.14, Arousal = 0.20) thresholds to indicate when the player has left a state of flow. From Figure 1 we can see that both arousal levels and engagement levels are mostly within the threshold levels during general game play. During death events participants leave the threshold levels indicating flow has been disrupted.

3. Discussion

Our goal was to coordinate “Task Engagement” and “Arousal-Valence” data to establish a flow model with the use of the Emotiv during video game play. We used the Emotiv to analyze difference in engagement levels, arousal levels, and valence levels between specific game events (general game play and death events). The primary results were: (a) significantly increased engagement levels were found during death events compared to general game play events; (b) significantly increased arousal levels were also found during death events when compared to general game play.

3.1. Applying thresholds to indices

Findings suggest that engagement (Beta / (Alpha + Theta) and arousal (((BetaF3 + BetaF4) / (AlphaF3 + AlphaF4)) indices can measure immersion levels. Higher levels of arousal and engagement were measured during death events and compared to general game play. Higher levels of engagement during death events may not suggest the user is more engaged or aroused when their character dies, but may reflect that they have entered a more stressful state which has increased their vigilance [19, 20]. Putting thresholds on player’s engagement levels based upon their results would help identify when players have entered a stressful state and identify when a death event has occurred. Establishing thresholds is a complicated task due to variability in the EEG data. Threshold levels are not fixed factors and will need to be adjusted as more player data is incorporated. However, using thresholds along with combining Task Engagement and Arousal-Valence we built a set of rules to indicate when the player is in a state of flow. 1) If engagement levels fall below the lower threshold, then the game is to become more complex; 2) If engagement level rises above the upper threshold, then the game is to become simpler; 3) If arousal level falls below the lower threshold, then the game play is to be more stimulating; and 4) If arousal level rises above the upper threshold, then game play is to become less arousing. These rules can be applied to any method or variation in the threshold levels.

3.2. Limitations and future directions

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 [16, 28]. Future
studies will be needed to expand these results into methodological approaches to quantifying videogame based EEG assessment in general and Emotiv–based EEG assessment of various games in particular.

4. Conclusion

We have presented findings from a study aimed at validating the use of the engagement index (Beta / (Alpha + Theta), arousal index (BetaF3 + BetaF4) / (AlphaF3 + AlphaF4), and the valence index (AlphaF4 / BetaF4) - (AlphaF3 / BetaF3) with the Emotiv EEG headset. We also aimed at coordinating “Task Engagement” data with “Arousal-Valence” data to establish a flow model. As our results show we were able to find significantly higher levels of engagement and arousal during death events when compared to general gameplay. Our finding suggests that combining engagement data with arousal data we can establish thresholds to indicate when a player has exited a state of flow. Applying rules to our model allows us to respond adequately to the changes in level of immersion the players are experiencing. It is believed that this model will allow for future use of the Emotiv for assessing the cognitive and emotional processing of the player.

References