A PSYCHOPHYSIOLOGICAL ASSESSMENT OF FEAR EXPERIENCE IN RESPONSE TO SOUND DURING COMPUTER VIDEO GAMEPLAY

Tom A. Garner\textsuperscript{1} and Mark N. Grimshaw\textsuperscript{2}

\textit{University of Aalborg}
\textit{School of Communication}
\textit{AALBORG, Denmark}
\textsuperscript{1}0045 71 58 73 38
\textsuperscript{2}0045 99 40 91 00

ABSTRACT
The potential value of a looping biometric feedback system as a key component of adaptive computer video games is significant. Psychophysiological measures are essential to the development of an automated emotion recognition program, capable of interpreting physiological data into models of affect and systematically altering the game environment in response. This article presents empirical data the analysis of which advocates electrodermal activity and electromyography as suitable physiological measures to work effectively within a computer video game-based biometric feedback loop, within which sound is the primary affective stimuli.

KEYWORDS
Psychophysiology, biofeedback, affective sound, adaptive gameplay

1. INTRODUCTION

The overarching problem that motivates this study is the insufficient capacity of computer software (specifically recreational computer video games [CVG]) to respond to the affective state of the user. Prior research has stated that this limitation significantly damages usability between human and computer (Picard, 2000). From a CVG perspective, the absence of an affect recognition system can limit: the effectiveness of social/emotional communication between players and virtual characters, the game’s capacity to respond to undesirable gameplay experiences such as boredom or frustration, the opportunity for the system to build an affective user-profile to automatically customise game experiences, and also, the potential to communicate emotions to other live players over a network.

Electrodermal activity (EDA) is characteristically related to the sympathetic nervous system (Nacke et al., 2009) and consequently, automated and excitatory processes (Poh et al., 2010). Yokota et al. (1962) connected EDA to emotional experience based upon a correlation between EDA and neural activity within the limbic system; an association that has been confirmed in later research via functional magnetic resonance imaging (Critchley et al., 2000). Relevant research has connected EDA to pathologic behaviour and stress (Fung et al., 2005), fear (Bradley et al., 2008) and disgust (Jackson et al., 2000) but, arguably, the most characteristic use of EDA is as a measure of general human arousal (Gilroy et al., 2012; Nacke & Mandryk, 2010). Research has suggested that EDA has the potential to measure changes in cognition and attention (Critchley et al., 2000) and, in conjunction with additional psychophysiological measures, may be capable of identifying discrete emotional states (Drachen et al., 2010). Electromyography (EMG) measures the voltage difference that contracts striated muscle tissue (Gilroy et al., 2012) and can be applied to various muscles around the body via either intramuscular (internal) or surface application. Studies of facial muscular activity have associated EMG to emotional valence, most typically by way of the corrugator supercilii (negative affective association) and the zygomaticus major (positive affect; see Lang et al., 1998; Larsen et al., 2003).
Academic literature has advocated physiological measures, such as EDA and EMG, for practical applications that include usability and user experience testing (Gualeni, 2012; Ravaja et al., 2006). Furthermore, these biometric measures have also featured in affect studies that employ sonic stimuli (Koelsch et al., 2008; Roy et al., 2008). Research utilising biometrics within a CVG context is becoming increasingly expansive, with forays into biometric-based adaptive difficulty systems (Ambinder, 2011), physiological measures for emotion-related CVG testing (Keeker et al., 2004) and even emotionally responsive game environments (Sakurazawa et al., 2004). The psychophysiological effects of computer game sound effects (excluding music and speech) have been underrepresented within this field of study despite a consensus that sound is a particularly evocative modality (Tajadura-Jiménez & Västfjäll, 2008).

The experimentation documented within this article takes its influence from the above research. EDA and EMG signal data is collected from two groups of participants; both playing a bespoke game level. The design of both the control and test game levels is identical, with the exception of digital signal processing (DSP) sound treatments that overlay particular sound events in the test group. DSP treatments are then compared to control datasets in a search for significant difference in arousal, corrugator supercilii activity and qualitative post-play feedback. As an exploratory study, it is hypothesised that both EDA and EMG measures will reliably reveal physiological changes in response to game sound stimuli. It is further hoped that at least some of the test sounds will generate significantly different datasets between groups. It is also expected that the physiological data will reflect subjective responses presented by participants during the debriefing. This article forms an essential element of a larger study, that aims to assess a comprehensive range of psychophysiological parameters within a CVG/sound/fear context and ultimately develop a new software biofeedback system that can accurately determine players’ emotional states and adapt the gameplay environment in real-time. It is further anticipated that such a system could be utilised beyond CVG applications with any interactive product, from cellular phones to automobiles.

2. METHODOLOGY

2.1 Bespoke Game Design

To enable effective comparison of the desired audio variables, a bespoke first-person perspective game level was developed, entitled The Carrier. This game places the player in the dark bowels of a sinking ship, with a race against time to reach the surface. The presence of a dangerous creature is alluded to via scripted animation sequences within the gameplay, and the intention is for the player to believe that they are being hunted. The level was produced primarily using the CryEngine 2 sandbox editor (CryTek, 2007) and the level designs incorporate a sequence of prescribed events, designed to subtly manipulate the player’s actions. Plausible physical barriers, disabling of the run and jump control functions and a linear progression of game scenes, restricts the player into following a more uniform direction and pacing. These constraints are complemented by the reduced visibility settings, which provide plausibly restricted vision and movement to encourage (rather than force) players to follow the desired path. Ambient atmospheres and sound events of indeterminate diegetic status positioned in the darkness further the perception of a larger, open world to add some credibility and realism to the game environment despite its linear design.

The opening scene of the level presents the premise and endeavours to create an initial sense of familiarity and security via recognisable architecture and everyday props. This atmosphere is juxtaposed against a dark and solitary environment to create a sense of unease from within the familiar. Subsequent scenes utilise conventional survival horror environments whilst implied supernatural elements and scenarios also draw heavily from archetypal horror themes. First-person perspective is retained but the customary FPS heads-up display and weapon wielding is omitted, giving the player no indication of avatar health or damage resistance, and also removing the traditional ordnance that increases player coping ability and diminishes vulnerability-related fears. The avatar has no explicit appearance, character or gender and is anchored into the gameplay via physics-generated audio (footsteps, rustling of vegetation, interactions with objects, etc.) and the avatar’s shadow. The player is required to navigate the level and complete basic puzzles to succeed. Unbeknownst to the player, their avatar cannot be killed or suffer damage to ensure that load/save elements are unnecessary and that no player will repeat any section of gameplay, thus further unifying the collective experiences of all participants.
2.2 Game Sound Design

All sounds within the test game were pre-treated in Cubase 5.1 (Steinberg, 2009) and separate sound files were generated for both variations of each key sound. For the purpose of this experiment, the seven modalities generated twelve key sounds (two sounds for each modality – to support the argument that if a DSP effect were to generate a significant difference, this would be observable when tested on two different sounds). Due to time limitations and gameplay restraints, signal/noise ratio and tempo parameters could only be tested once per game type. Two variations of each sound were developed as contrasting extremes of each modality, producing a total database of 24 files per game. Figure 1 outlines the use of sound employed throughout both test levels.

<table>
<thead>
<tr>
<th>Sound Name</th>
<th>DSP modality</th>
<th>Control (group A)</th>
<th>Variant (group B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diegetic Music</td>
<td>Distortion</td>
<td>No additional DSP</td>
<td>Frequency distortion</td>
</tr>
<tr>
<td>Ship Voice</td>
<td>Distortion</td>
<td>No additional DSP</td>
<td>Frequency distortion</td>
</tr>
<tr>
<td>Heavy Breath</td>
<td>Localisation</td>
<td>Centralised</td>
<td>Left to right sweep</td>
</tr>
<tr>
<td>Monster Scream</td>
<td>Localisation</td>
<td>Centralised</td>
<td>Full left pan</td>
</tr>
<tr>
<td>Woman Screams</td>
<td>Pitch</td>
<td>No additional DSP</td>
<td>300 cent pitch raise</td>
</tr>
<tr>
<td>Ship Groans</td>
<td>Pitch</td>
<td>No additional DSP</td>
<td>500 cent pitch drop</td>
</tr>
<tr>
<td>Chamber banging</td>
<td>Attack</td>
<td>2 second linear fade-in</td>
<td>0 second attack</td>
</tr>
<tr>
<td>Monster Growl</td>
<td>Attack</td>
<td>1 second linear fade-in</td>
<td>0 second attack</td>
</tr>
<tr>
<td>Bulkhead Slams</td>
<td>Tempo</td>
<td>20 BPM</td>
<td>30 BPM</td>
</tr>
<tr>
<td>Engine Noise</td>
<td>Signal/noise ratio</td>
<td>No noise present</td>
<td>Noise present</td>
</tr>
<tr>
<td>Man Screaming</td>
<td>Sharpness</td>
<td>No additional DSP</td>
<td>12 dB gain @ 1.7kHz</td>
</tr>
<tr>
<td>Man Weeping</td>
<td>Sharpness</td>
<td>No additional DSP</td>
<td>12 dB gain @ 5kHz</td>
</tr>
</tbody>
</table>

2.3 Testing Environment and Equipment

The game level ran on a bespoke 64-bit PC with Windows Vista Home Premium (Service Pack 2) operating system, AMD Phenom 2 X4 955 (3.2GHz) quad core processor, 8GB RAM, ATI Radeon 4850 (1.5GB) GPU. Peripheral specification includes LG 22” LCD Monitor (supporting 1920x1080 output resolution), Microsoft Wireless Desktop 3000 mouse and keyboard, Asus Xonar 7.1 sound card and Triton AX Pro 7.1 headphones. Fraps (Beepa, 2007) screen capture software created video records of all gameplay and biometric data was collected using a Biopac MP30 data acquisition unit and Biopac Student Lab Pro v3.6.7 (Biopac, 2001) interface software. Experimentation was carried out in a small studio space, providing only artificial light and attenuation of outside environment noise.

2.4 Participants, Procedure and Ethics

10 participants (9 Male, 1 Female, ages 18-27) formed two groups of 5, control and test (DSP treatment). All participants rated their prior experience and gaming confidence as moderate or high and stated familiarity with FPS type games and PC standard gaming controls. Experience in survival-horror games revealed some variation, with self-report ratings ranging from 1 to 10 (1-10 scale) with a mean score of 4.7. Participants were informed that the research aim was to explore the emotional potential of sound within a computer video game context. Each individual was also made aware of strobe lighting, visual images (that may be perceived as frightening or upsetting) and the full biometric data collection procedure. A satisfactory EMG/EDA baseline was acquired before beginning testing and synchronisation of gameplay with both the biometric and video recordings was achieved by mapping the respective start and stop actions to the same key, allowing the participant to synchronise the entire data collection process with little difficulty. The test debriefing assessed perceived game difficulty, overall intensity of experience, immersion ratings and disruption caused by the biometric sensors.
Participants were also asked to watch the video of their gameplay performance and provide a voice-over commentary, focussing upon their emotional state and identifying discrete emotions, the intention being that participants would re-experience the affective states felt during the game and be able to more accurately describe their emotions in reference to specific game events.

2.6 Data collection

EDA and EMG hardware was configured to automatically synchronise with the game engine timestamp, allowing significant biometric readings to be accurately matched with their corresponding chronological point of gameplay. An event logging system was utilised to identify overall completion time. EDA data was collected from the right index and middle fingers of each participant by way of a SS57L Biopac EDA sensor lead and isotonic electrode gel. BSL shielded SS2LB leads connected to trimmed, disposable EL501 electrodes were utilised to collect facial EMG data. Existing research warns that precise positioning of EMG electrodes is a difficult task (Huang et al., 2004), therefore utmost care was taken to apply the hardware to each participant. A light abrasive treatment was applied to the skin in and around the areas on which the sensors would be placed to reduce electrode-skin impedance (Hermens et al., 2000). Electrodes were then applied across the midline of the muscle (De Luca, 1997) and surgical tape was used to reduce motion artefacts (Huang et al., 2004). The validity of the zygomaticus major as a measure of positive affect has been questioned in response to an assertion that conscious social communication may be the primary way in which this muscle is controlled (Russell et al., 2003). Larsen et al. (2003) however, suggest that the corrugator supercilii is capable of representing both positive and negative valence (suppressed corrugator activity suggesting positive affect). In response to this information, EMG data was collected solely from the corrugator muscle.

3. RESULTS

Data obtained from both EDA and EMG acquisition was extracted in 5.00 second epochs. The baseline epoch for each participant was collected between 30.00 and 35.00 seconds of the 1 minute rest period before the game began. This was to allow time for the signal to stabilise test-start anxiety (white coat hypertension). Data relevant to each test sound was extracted as 5.00 second epochs in synchronisation with the sound’s onset. Integrated signal analysis tools within BSL Pro v.3.6.7 (Biopac, 2001) revealed descriptive statistical data across each epoch. Both EDA and EMG data was analysed for minimal and maximum peaks, mean, area and slope. Subsequent statistical analysis was carried out using PASW statistics (v.18, IBM, 2009) to calculate the difference between the baseline and each test epoch to generate a differential dataset. The descriptive statistics of the differential dataset were then analysed via a statistical t-test for independent samples, to search for statistically significant difference between the means of the control and treatment groups for each test sound. Results of the t-test revealed significant difference between groups of the muffled scream (sharpness treatment) sound in the EDA-mean (t=-2.377, p=.045), EDA-max (t=-2.357, p=.046) and EDA-min (t=-2.457, p=.039) dependent variables and was the only significant difference in the EDA output measures. Of the EMG outputs, only EMG-max revealed a significant difference between groups, at the muffled whimper (also sharpness treatment) sound (t=2.669, p=.028). Each individual sound (both treated and untreated) was separately compared but no statistical significance between any of the sounds for any measure of EDA or EMG was revealed. An overview of the descriptive statistics suggests that overly high variance between players, present even when testing the differential data, is the likely cause of an absence of significant difference between sounds.
Figure 1 presents a visual overview of the EMG and EDA signal data from one subject. Overlaying biometric output with event data reveals that the majority of EDA/EMG spikes and peaks can be attributed to specific events within the game. However, not all of these events are exclusively tied to a sound and, quite often, also relate to a visual stimulus. Visual analysis of the biometric signals also reveals significantly raised skin conductance levels during the first 60-80 seconds of gameplay, and a steadily increasing rate of EMG activity from commencing gameplay to completing the level. These trends are generally representative of all players’ test results.
Non-parametric analysis compared the control and test groups via the Mann-Whitney U test for two independent samples to assess if any significant variation existed between the groups with regards to subjective intensity, frustration and difficulty ratings. Results revealed no statistical significance between groups for any of the three dependent variables. Descriptive statistics and the Chi-Square test was administered, including all 10 players’ results, to assess comfort of biometric sensors, perceived difficulty of the game itself, subjective frustration levels and the extent to which the sensors and wires obstructed the flow of gameplay and sense of immersion.

The results of the Chi-Square test reveal statistical significance for all measures and the descriptive means reveal a relatively high level of comfort and low disruption during gameplay alongside moderately high intensity ratings, low difficulty and very low frustration levels. Information collected from qualitative discussion with participants revealed that context/situation heavily influenced how they felt about a sound. For example, many commented that the consecutive slams at the end of the level created suspense and fear by signifying urgency and a time-limit to reach the level end. Participants also commented that the alien roar sound was very intense and discomforting. Participants did not comment directly upon the fast tempo of the slam sounds or the immediate attack and high volume of the alien roar sound, suggesting limited conscious awareness of the quantitative sound techniques being employed. Figures 2 and 3 present line graph representations of particular statistical measures of both EEG and EDA, revealing the difference between sound treatment groups and individual sounds.

4. DISCUSSION

As an exploratory experiment, several concerns were foreseeable from the onset whilst others were revealed throughout the course of the study. The overarching themes, atmospheres and stylisations of a computer video game are a likely source of variation upon the perception of individual sounds and compound sound events/ambiences, the rationale being that such factors form a contextualising framework against which game stimuli are appraised. Future work will consequently incorporate two contrasting game levels, providing an opportunity to discover whether acoustic parameters could reliably alter affective states across varied contextual environments. The Biopac EDA/EMG hardware and accompanying software proved to be a powerful yet accessible solution. The integrated signal analysis features provide highly usable descriptive statistics upon which further analysis can be performed. Connection between players and sensors proved consistent, and the positive reviews from players concerning comfort of use and lack of flow/immersion disruption further advocates both the Biopac system and EDA/EMG biometrics in general as effective data acquisition tools within CVG research.

With regards to the game level design and control interface; observation of the participants’ movements, in conjunction with debrief questionnaire responses, revealed a generally high level of usability, with most players able to move effectively throughout the level. The lack of a heads-up display (HUD), text-based instruction and mini-map navigation was regarded by some players as initially confusing, but not a source of frustration.
Feedback comments also suggested that the absence of an extra-diegetic HUD improved immersion and that undetermined player-health feedback increased tension. Players also commented that an absence of weapons signified no combat and they quickly realised that they could not die, a realisation that quickly reduced fear intensity by way of removing the threat, both within the diegetic narrative (no threat of avatar damage/death) and as an extra-diegetic gameplay tension (no threat of having to repeat part of the level). With reference to EDA, it could be asserted that ecological validity and laboratory-based bias reduced the capacity to extract meaningful data. It was noted that various aspects of the testing environment (continuous presence of researcher, white-coat syndrome from biometric hardware application and, unfamiliar environment, control hardware, action-button configuration, audio setup and computer monitor / graphics setup) were likely to raise anxiety levels prior to testing. The particular durations and placement of data extraction epochs present additional potential for erroneous effects, the most notable problem being the limited time between starting the biometric recording and starting the game (within which the baseline epoch is collected). Observed EDA trends strongly suggest a consistently occurring peak in activity, lasting for up to 90 seconds after beginning the recording and a long decay extending over several minutes before a lower-level plateau is eventually reached. This suggests that a more substantial ‘settling in period’ is required between beginning the recording and the commencement of gameplay.

Whilst the lack of statistically significant difference between the test and control groups could be disheartening, it must be remembered that the primary function of the experiment was to assess the potential of EDA/EMG within an audio-centric survival horror game context, uncover potential methodological problems and present possible solutions for use in future study. Furthermore, the correlation between quantitative biometric data and qualitative debrief responses suggests that data acquisition was notably successful and that the issues are more likely to be associated with the testing environment, equipment and procedure.

5. CONCLUSIONS

Although the results obtained from this study do not fully support the hypothesis that quantifiable acoustic parameter changes can generate significantly different psychophysiological response datasets, they do support the use of biometrics within such experimental scenarios and present a foundation from which to build in further study. This article strongly advocates a systematic approach to game sound testing in future work, to allow for the various erroneous factors to be addressed. Ecological validity is a substantial concern and future testing will attempt to recreate a familiar and comfortable playing environment. Researcher presence should arguably be minimised and players should not be informed when recording has begun to reduce white-coat syndrome. Future testing will also integrate video tutorials (reducing live researcher presence and adding uniformity to the participant instruction process) and present the player with an interactive tutorial level, allowing them to become familiarised with the interface, control mapping, etc. and also providing time for stress factors associated with initial exposure to the testing environment to subside.

Direct experience with the psychophysiological measures of electromyography and electrodermal activity supported many of the assertions referenced within the introduction of this article. EDA proved to be a reliable indicator of arousal and in context, this could be restated as a measure of gameplay-related fear, anxiety or stress. The EDA equipment was, as expected, easy to apply and operate (Boucsein, 1992; Nacke & Mandryk, 2010) with minimal intrusion for the player (Lorber, 2004). Motor activity (Roy et al., 2008) proved not to be a source of erroneous variation and the temporal resolution, traditionally described as slow (Kivikangas et al., 2010) proved adequate for accurate synchronisation to in-game stimuli. EMG confirmed expectations for high temporal resolution and sensitivity (Bolls et al., 2001) but did not demonstrate itself to be a reliable indicator of negative valence, contradicting much previous research (Harman-Jones & Allen, 2001; Kallinen & Ravaja, 2007). This outcome does, however, support the notion that recreational fear is an ambivalent experience (Svendsen, 2008: pp. 75-76). In terms of fear experience, the data obtained from these three experiments strongly supports the concept of a complex, embodied system that is susceptible to reflexive shock, slow-building apprehensive/suspenseful terror and more subtle variations between the two. Despite strong effort to control erroneous variables during testing, between-participant differences remained noteworthy, supporting the notion of interpersonal affective influences such as gender, culture and personality (Hamann & Canli, 2004).
With regards to the secondary hypothesis of this article (assessing the emotional/affective potential of quantifiable acoustic parameters) the obtained data does reveal both resonance and dissonance when considered alongside existing research. Parameters such as immediate attack (Moncrieve et al., 2001), increasing tempo (Alves & Roque, 2009), low pitch (Parker & Heerema, 2007), and unclear localisation (Breinbjerg, 2005) are supported by the results obtained from these experiments, but not conclusively. However, sharpness (Cho et al., 2001) does, based upon obtained data, appear to be a significantly reliable approach to modulating both EDA and EMG. The notion that some sounds may have the capacity to universally evoke a particular emotional state by way of underlying evolutionary factors (Parker & Heerema, 2007) remains uncertain due to the presence of contradictory data (for example, participants producing a limited or negligible affective response to sounds specifically designed to evoke evolution-based fear responses). The assertion that sound can be processed pre-attentively (Alho & Sinervo, 1997) is supported by the notably faster response times of EMG (and, in some cases, even EDA) to stimuli than the real-time qualitative responses of players. Indirectly, the lack of clear patterns within the data could be perceived as supportive of the concepts that imply auditory processing as a complex matrix of variables. This includes the discrete modes of listening (Chion, 1994), embodied cognition factors (Wilson, 2002), attention filtering (Ekman, 2009), multi-modal effects (Adams et al., 2002) and sonification data (Grimshaw, 2007; Schafer, 1994). As a result, data obtained from this experiment indirectly supports a hypothetical framework of auditory processing that incorporates such concepts.

REFERENCES


