Analysis of Player Stress Levels Using the Affectiva Q Sensor

Gabriel Jose G. Vitug
Ateneo de Manila University
gjgv124@gmail.com

James Alexander D. Toplis
Ateneo de Manila University
james_123456789_james@yahoo.com

Regina Ira Antonette M. Geli
Ateneo de Manila University
riamgeli@gmail.com

ABSTRACT
Graphics and sounds within video games have evolved throughout the years to create a more immersive environment for players. Video games are engineered to keep the players in a continuous state of arousal through the combination of graphics and sounds. Through the in-game events, players remain in a state of arousal, keeping the immersed in the game environment.

Our study is focused on finding out the effects of arousal-inducing events in a horror video game on players’ stress levels. By having our participants play Slender while wearing the Affectiva Q Sensor, we wanted to determine which in-game events triggered arousal and if the presence of game sounds affected their stress levels.

Our findings showed that these three in-game events triggered arousal: seeing Slenderman (the antagonist in the game), seeing static on the screen, and finding a page. However, our findings suggest that game sounds did not have significant effects.

Categories and Subject Descriptors
G.3 [Probability and Statistics]: Experimental design, Reliability and life testing, Statistical computing, Statistical software; D.4.8 [Performance]: Measurements, Operational analysis; D.m [Miscellaneous]: Software psychology

General Terms

Keywords
Arousal, stress, emotion, immersion, electrodermal activity. Slender, Q Sensor

1. INTRODUCTION
One of the elements considered in the design of video games is how these games are able to evoke emotion within the player. These emotions are tied to the presence of elements such as sound and visuals [2].

Video games are engineered to manage the flow of arousal. Good games sustain arousal by including quickly changing features synchronized with the natural flow of human attention [10]. Developers strive to improve games such that visual cues and audio rhythm, for example, are able to heavily influence the retention of the player’s attention to prompt arousal, and subsequently, stress [10].

To this end, games have grown in sophistication. In terms of visual elements, games began with only simple 8-bit pixels for graphics but eventually move to 3D with the arrival of games such as Battlezone and 3D Monster Maze [9]. A better example of how visuals aid in the player’s experience of the game is the Oculus Rift, which allows players to enter the virtual world of the game as their player character [7].

Music has likewise seen the same evolution. Just as graphics started with 8-bit pixels, video game audio began with a simple hollow ringing sound from the video game Pong. Technology in audio has risen significantly since then. MP3 Compression and the development of the CD-ROM and DVD are only some of the examples of audio equipment that has helped with the evolution of music and sound [12].

Partnered with modern graphics, sound helps to create better games which evoke arousal and stress in players. Sound will be the main object of study in this experiment and will hopefully give light to its effects on stress and arousal.

The goal of this study is to determine how stress levels in participants playing a horror video game are affected by arousal-inducing video game stimuli by answering the following research questions:
1. What game events trigger arousal?
2. What effect does the presence of sound have on players as compared to players who played the game on mute?

2. DISCUSSION OF KEY CONCEPTS
Arousal and stress are closely related but not interchangeable. Stress is a broad term that can take on a myriad of meanings depending on the individual using them. Thus, it is necessary to define arousal and stress within the context of the paper and how they relate to each other. Electrodermal Activity (EDA) as the measurement of stress must also be defined.

2.1 Arousal
Arousal means a constantly changing state of readiness for action that is an engine for thinking and behavior, not necessarily excitement associated with exercise or erotic activity. Arousal is caused by specific occurrences from the environment that engages people to approach or avoid these events. It is rarely triggered by the person experiencing it [10].

Triggers of arousal include sudden movement, audio factors, interaction with other players, etc. These features are engineered to be conspicuous and salient so that, even though they are virtual, they are unconsciously and involuntarily engaging [10]. Unlike emotions such as happiness, sadness, etc., arousal only lasts from seconds to minutes and may spike repetitively. What does this imply? It implies that good timing and placement of the aforementioned elements of gameplay play a serious role in emotional engineering. [10].

2.2 Stress
Stress is described as “any influence which disturbs the natural equilibrium of the body.” [1] Arousal triggers stress [1]. It can
increase or decrease a player’s performance which means that a player will perform better under pressure if he or she has better stress management skills [1]. In a study researching the effects of stress and arousal on memory and performance, it was discussed that stress is the effect when arousal triggers the fight-or-flight syndrome by triggering the body’s adrenaline response. [3].

Stress is more clearly defined by a study by Smith and Sulsky. A study by [11] more clearly defines stress. Firstly, Smith and Sulsky define stress with the Simulus Definition, which states that stress is defined in terms of the events, or stressors, which garner a stressful response. This definition focuses on external factors that are applied to induce a stress. The Response Definition refers to the reaction the subject has towards the stressor. This response can take on multiple forms, from immediate, reflexive actions to more long-term actions [11]. This response is always internal and psychological, but may include physical manifestations such as trembling hands or goosebumps. One must also consider that a person’s response to a stressor may change and diminish over time and that certain stressors may not be seen as threatening to certain people [11].

This Stimulus-Response Definition describes stress as resulting from interactions between stressors and individual responses. Because of its broader and more complex focus, this definition is able to accommodate the influence of other factors such as a person’s appraisal of how stressful certain stressors are and how external factors affect their individual responses [11].

A person perceives a stressor as threatening or stressful through a psychological appraisal of the stressor that determines its severity [11]. A sheltered and weak person for instance, would consider walking down a dark alley much more threatening and stressful than, say, a man hardened from the streets. The term stress moderators are also used to describe outside factors such as location, or the presence of certain people, that can alter an individual’s response [11].

2.3 Electrodermal Activity
The measurement of electrodermal activity (EDA) allows people to create inferences about a player’s stress levels [5]. EDA is described as the measurement electrical conductivity of the skin. A measure on the relationship between stress and EDA done by Drachen [4] shows that there is indeed a correlation between the two. The study shows that high EDA values are associated with high levels of stress, and that a high level of stress can be “indicative of a high level of challenge, frustration, and/or excitement.” EDA does not measure arousal [4]. However, as explained, stress is triggered by arousal so measuring stress will give the researchers insight on how arousal was utilized by the elements of the video game.

3. EXPERIMENTATION
The goal of this study is to determine how stress levels in participants playing a horror video game are affected by arousal-inducing video game stimuli. The arousal levels of the participants were measured with the Affectiva Q Sensor. The control group of participants played a horror game with full audio while the experimental group played without any sounds or music. At the end of each test, the participants were asked to gauge their stress levels with a questionnaire. It is believed that the presence of ambient sounds and music will garner a greater arousal response from the test subjects. Therefore, the hypothesis is: the arousal levels of the participants will be greater in the control group than in the experimental group.

3.1 Participants
The target population of testers will include the following characteristics:
- Student of the Ateneo de Manila University
- 17-22 years old
- Is familiar with the navigation of video games of any platform.
- Has not played or seen gameplay footage of Slender

An initial number of ten participants were tested, but the data of three of the participants was removed and replaced by data collected from three new participants. Of the 13 total participants, two were female and 11 were male.

3.2 Design
Each participant was assigned to either play the game with or without ambient sounds and music. While each participant was randomly placed in one of the two testing groups, each group was balanced to have five participants.

3.3 Materials
The following items listed are the primary materials needed in the experiment:

Q Sensor - The Q Sensor (see Figure 1) is used to track increases and decreases in arousal, stress, or excitement within the context of a real world setting wherein the participant is not stuck in a lab or using a computer. First designed for the benefit of more effective communication with sufferers of autism, the device is used to detect and record physiological stress and excitement levels by measuring the electrodermal activity (EDA). The device sends readings which can be read by the included software Q Live. The Q Sensor is simple to install and connect to the laptop, only needing a Bluetooth connection. The device, worn on a wristband in the same manner as one might wear a watch, is designed to be small and comfortable, decreasing the distraction a lab or computer monitor might evoke, making data collection easy. [6].

Figure 1. Affectiva Q Sensor unit

Slender - Slender is a first-person game. The player’s character explores a wooded area collecting eight pages scattered throughout the environment. The eight pages contain scribbled messages or drawings (see Figure 3), all of which pertain to Slenderman (see Figure 2), the antagonist of the game. The player’s character dies if s/he encounters Slenderman. Atmospheric hints such as static on the screen (see Figure 4),
heightening volume of music, and increased frequency of eerie thumping sounds inform the player that Slenderman is near. Slender was chosen as the game for the experiment because the significant in-game events were easily recognizable (seeing Slenderman, seeing static, finding a page).

3.4 Procedure

The testing pool was divided into two groups. One group played the game Slender with sound, while the other group played the game on mute.

Only one test subject played at a time. None of the players in the testing pool had played Slender yet.

Each run of the game continued until the end of 20 minutes or until the player’s character died.

During the game play, the participant wore a Q Sensor. EDA readings from the Q Sensor were recorded through the Q Live.

Since one of the objectives was to pinpoint the events which garner emotional or mental responses, three events were chosen that had the highest probability of causing a response: encountering Slenderman, static on the screen, and finding pages. Whenever a participant encountered these events while playing, a time marker would be place to indicate the approximate beginning of an event.

We also used Fraps, a program which takes a video of activities running on the computer screen, to record the monitor while the participant is playing.

A questionnaire which asked about a self-assessment of stress and comments about the gameplay experience was given to the participant after they were done.

The set-up can be described as such: the player plays the Slender game on a computer in an isolated environment. Meaning s/he is free from distractions such as unnecessary noise. The player is wearing the Q Sensor during this time. The game being played is also being recorded by Fraps.

On a separate computer, we observed the data being fed to the Q Live software from the Q Sensor. During each test, several time markers were placed to indicate that an in-game event has occurred.

Using the Q Live software, we recorded their emotional states in conjunction with the current game events. The Q Live shows a line graph showing the subject’s arousal level in real time, indicated by time stamps. The line graph will increase in level as the player feels more stress. The Q Live graph is then recorded and saved. Each spike in the graph with a time marker was noted down and the time this happened was compared to the time of the recorded gameplay (recorded via Fraps) to compare what event evoked such a spike in stress level.

We annotated the data gathered according to the game events or player activity. Then, we averaged the readings for each event type to create numerical representation of the emotional response stemming from a certain action or event. We used the data analysis function of Microsoft Excel 2010 to statistically analyze the data from the participants to create a machine model that can show how engaged players are while playing a game.

Figure 5 shows a flowchart of our procedure.
4. DISCUSSION

4.1 Analysis

The time markers in each of the participants’ data do not give the entire EDA reading for each event since the markers only showed the approximate beginning of each in-game event these events lasted several seconds (see Figure 6). In order to get the complete reading of each event and better segment the data, the time stamp of each marker was used to synchronize each event along with its duration using the gameplay video recorded with Fraps. Once an event was synchronized with the video, its duration is marked on the line graph using a ranged marker along with a label of what kind of event occurred (see Figure 7).
After segmenting each reading, the line graphs and markers were converted to two .csv files each through the Q program. These .csv files can be opened by using Excel in Microsoft Office or any similar workbook programs. One .csv file contains every EDA reading of the participant along with the corresponding timestamp and the other .csv file contains the time range of each event (see Figure 8). Using these time ranges, the corresponding EDA readings were gathered from the first .csv file.

Each event’s readings were averaged in order to get the general EDA value during the events duration. The general EDA values of events of the same type were grouped together and averaged to get the participants general EDA values for certain events.

![Table on the left: Events and their time range, table on the right: EDA readings and their corresponding time. In this example, the EDA readings of the event "Slender Static" from 11:12:53 AM to 11:12:56 AM are being gathered.](image)

The EDA values of each participant are listed below (note that the values under the Slender, Static, and Page event columns are the amount increased from the Baseline value):

<table>
<thead>
<tr>
<th>Participant</th>
<th>Slender</th>
<th>Static</th>
<th>Page</th>
<th>Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.483</td>
<td>1.803</td>
<td>1.012</td>
<td>0.498</td>
</tr>
<tr>
<td>2</td>
<td>1.159</td>
<td>0.544</td>
<td>0.559</td>
<td>0.610</td>
</tr>
<tr>
<td>3</td>
<td>2.554</td>
<td>2.113</td>
<td>0.889</td>
<td>0.388</td>
</tr>
<tr>
<td>4</td>
<td>2.147</td>
<td>1.735</td>
<td>0.352</td>
<td>2.176</td>
</tr>
<tr>
<td>5</td>
<td>2.269</td>
<td>1.683</td>
<td>0.947</td>
<td>0.239</td>
</tr>
<tr>
<td>Average</td>
<td>1.922</td>
<td>1.575</td>
<td>0.752</td>
<td>0.782</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.519</td>
<td>0.537</td>
<td>0.254</td>
<td>0.708</td>
</tr>
</tbody>
</table>

![Table 2. Average results for without sound](image)

<table>
<thead>
<tr>
<th>Participant</th>
<th>Slender</th>
<th>Static</th>
<th>Page</th>
<th>Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.390</td>
<td>3.612</td>
<td>2.717</td>
<td>2.509</td>
</tr>
<tr>
<td>2</td>
<td>0.703</td>
<td>0.596</td>
<td>0.433</td>
<td>0.558</td>
</tr>
<tr>
<td>3</td>
<td>2.032</td>
<td>1.613</td>
<td>1.175</td>
<td>0.455</td>
</tr>
<tr>
<td>4</td>
<td>1.437</td>
<td>0.704</td>
<td>0.287</td>
<td>0.851</td>
</tr>
<tr>
<td>5</td>
<td>1.374</td>
<td>0.385</td>
<td>0.374</td>
<td>1.706</td>
</tr>
<tr>
<td>Average</td>
<td>1.787</td>
<td>1.382</td>
<td>0.997</td>
<td>1.216</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.905</td>
<td>1.191</td>
<td>0.917</td>
<td>0.782</td>
</tr>
</tbody>
</table>

(Note: EDA readings are recorded in microsiemens)

The values under the Slender, Static, and Page event columns are the results after the raw values were subtracted by the values under
the Baseline column. This was done so that the changes in EDA increase per event could be more accurate.

We use a two-factor ANOVA without replication on each table (With Sound and Without Sound) to check if the general EDA values were statistically significant. The t-test, assuming equal variances, was also used to compare the corresponding events in both tables (e.g. comparing the Static column in the With Sound table to the one in Without Sound).

### 4.2 Results

We attempted to determine whether the presence of music had an effect on the EDA responses of the participants.

First, the EDA values in the columns of the tables were tested in order to find out if the values were statistically significant. This was done using the two-factor ANOVA without replication. For the With Sound table, the test showed that the readings were statistically significant ($F(3,2) = 13.357; p = 0.003$). Testing the Without Sound table readings also showed that the readings were also statistically significant ($F(3,2) = 10.153; p = 0.006$).

Upon initial observation of the two tables containing the EDA readings, those who played the game with sound had slightly different EDA levels than those who played without sound. Those who played with sound had a higher EDA reading when encountering static on the screen (1.575 microsiemens $\mu$S on average) and Slenderman (1.922 $\mu$S on average) than those who played without music (1.382 $\mu$S on average for static on the screen, 1.787 $\mu$S for Slenderman). However, the EDA reading of those who played without sound for the page finding event (0.997 $\mu$S on average) was higher than those who played with sound (0.752 $\mu$S on average). A possible explanation to this would be that those who were playing with sound became more relaxed when finding pages (which is the objective of the game) since finding a page meant a short break from the tension of searching around for more pages. Those who were playing without sound might have relaxed a bit but were still alert since they did not have audio queues of what could possibly happen while retrieving a page.

To test whether the differences between the EDA values of the two tables were actually significant, the columns in the With Sound were compared to the columns in the Without Sound table using a T-test assuming Equal Variances. The Slender column in the With Sound table was paired with the Slender column from the Without Sound table for the test. The same pairing was done for the Page and Static columns.

During the Slender event the participants in the control (with sound) condition had, on average, higher EDA values (1.922 $\mu$S) than those in the experimental (without sound) condition (1.787 $\mu$S). However, the results were not statistically significant ($t(8) = 0.259$; two-tailed $p = 0.802$). For the Static event, the participants in the control condition also had, on average, higher EDA values (1.575 $\mu$S) than those in the experimental condition (1.382 $\mu$S). However, the results were also not statistically significant ($t(8) = 0.296$; two-tailed $p = 0.775$). Finally, for the Page event, the participants in the control condition had, on average, lower EDA values (0.752 $\mu$S) than those in the experimental condition (0.997 $\mu$S). However, the results were also not statistically significant ($t(8) = -0.516$; two-tailed $p = 0.620$). These results indicate that while there is a slight difference in results between the tables, none of them can be considered statistically significant meaning there is no difference major difference in the results of both tables.

One of the objectives was to prove that there would be a difference in EDA values between those who played Slenderman with sound and without sound. The results show that sound does not make a major difference in the EDA values. Our initial prediction was that there would be a significant difference between the EDA values of the two groups. Based on this, however, we can say that the level of stress between playing a game with and without sound is similar. Based on the results of the ANOVA tests, we can say that these events are actual shifts in the participants’ EDA readings which mean that they do garner an emotional or mental response. However, the T-test shows that the responses from both groups are similar. In the questionnaire we gave out after each play through of the game, we asked the participants how stressed they felt while playing. Below is the tally:

<table>
<thead>
<tr>
<th>With Sound</th>
<th>Without Sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
</tr>
</tbody>
</table>

It can be observed that the participants’ own rating of stress while playing the game only differs slightly between the two groups, a rating of 8, on average, for the With Sound group, and a rating of 7.8, on average, for the Without Sound group. The general comments about the game from the two groups were that Slenderman always caught them off guard when he appeared and that the game in general was scary. The ratings the participants gave do support the EDA readings gathered. The stress caused by playing Slender is the same, whether played with or without sound. This indicates that playing with or without sound induces the same level of arousal.

### 5. CONCLUSION

Through our experiment, we wanted to determine which game events triggered arousal in our participants. After putting our results through the two-factor ANOVA with replication, we can conclude that seeing Slenderman (Slender), seeing static on the screen (Static), and finding a page (Page) are in-game events which trigger arousal.

We also wanted to determine if the presence of sound had an effect on the players’ stress levels. The T-test assuming Equal Variances results showed that there was no statistical difference between those who played with sound and those who played without sound. We therefore conclude that the effects of sound on players’ stress levels were not detectible in this instance.

### 6. RECOMMENDATIONS

We would recommend that further studies should be done using a bigger sample size. Our study might have been limited due to the small sample size and it is possible that this may have affected the outcome of our experiment. We also recommend using games from other genres in order to find out if different genres can possibly result in different levels of arousal.
7. ACKNOWLEDGMENTS
The group would like to extend special thanks to Ma. Mercedes Rodrigo for guiding us throughout this endeavor. We also thank our testers who agreed to be terrified in the interests of science.

8. REFERENCES