# Creative Technologies Project:
## Analysing the use of Biometric Technology in the Field of Games User Research

**Author Name**
pearce2.jennings@live.uwe.ac.uk
Supervisor: Mark Palmer

**Department of Computer Science and Creative Technology**
University of the West of England
Coldharbour Lane
Bristol BS16 1QY

**Research Report**

**Abstract**

This report summarizes research that looks at the use of Biometric Technology in the field of Games User Research (GUR). This research will investigate how this technology is used to gain insight into player experience, primarily in relation to the collection and utilization of heart rate (HR) and galvanic skin response (GSR) data. A growing method adopted within the field is to take this data and cross reference it with data gathered from industry standard techniques like interviews and questionnaires, in order to recognize trends and then infer things about the players' experiences. The goal of the project will be to try and replicate this process, identify any limitations that the methodology has and ideally produce results that show a consistent correlation between different data sources.

**Keywords**: Games User Research, Heart rate, Galvanic skin response, Biometric Technology

## 1. Introduction

The field of Games User Research comprises of a collection of methods that allow designers to "bring their creations closer to the initial vision of the player experience" (Nacke, 2015, p.63). These methods involve observing and evaluating player experience in an attempt to improve the quality of a game, which is increasingly carried out these days through the use of Biometric Technology. This technology can be used to gather data during play sessions, that can then be used in conjuction with psycholigical principles to help iterate design. Physiological measurements can be used to differentiate between human emotions such as anger, frustration and excitement (Ekman *et al* 1983), which can give insight to developers about the players experience. Physiological evaluation is rapidly becoming a standard tool in user testing but it has limitations, as Biag *et al* (2019, p.2) states "there is no direct relationship between psychological phenomena and physiological processes, only a suggestive one", which is worsened by the fact that "interaction also changes from person to person".

Given that it is difficult to draw a definitive link between the two, it is crucial that tests involving multiple different users are conducted when trying to infer any sort of relationship between physiological data and the players experience. Various methods of testing have been carried out within the field of GUR and many have yielded inconclusive results (Mirza-Babaei *et al*, 2013), with other studies that cross reference different data sources having the most success (Drachen *et al*, 2010). This project will therefore cross reference three different sources of data; two different types of biometric data collected from user testing (HR and GSR), and quantitative data from questionaires.

Even though a lot of self reported feedback in the field of GUR is often qualititative, for the purpose of this project, the data collected will be quanititative and will provide answers that can be assigned numerical value, to try and draw a more objective correlation between the different sources of data. The biometric data will be collected via hand and wrist sensors connected to an Arduino, and the projects implementation will take the form of a system in Unity that works alongside Arduino's IDE. This system will gather HR and GSR data during testing and will then allow for the input of the questionaire results. The system will then cross reference these data sources (using techniques that will be covered in this reports findings) in order to produce a set of results for each participant. The way that these results will be correlated and visualised will also be covered in this reports findings. Once all the data is collected and analysed by the system, the results from each participant can then be compared with one another and ideally a consistent pattern will be present.

The main points that this research will address:

- How the questionaires will be carried out
- Collecting HR/GSR data, the logistics of how that data is analysed and what it can infer
- Issues involved with keeping the data reliable when testing with these methods
- What techniques the system will use to cross reference the results
- How these results will be visualised

## 2. Research methods

Gathering GSR/HR data is a process that has undergone a great deal of iteration in multiple different industries especially in the context of GUR and it has been optimised a lot in recent years, so the methodology of how to collect and utilize such data is well established.

As a result, qualitative secondary research is the method of research that will be used in this report to provide insight on the issues posed by the project. Information will be sourced from academic articles, conferences, previous studies and books.

## 3. Research findings

**Choosing a game for testing**

Using narrative driven games in testing has been found to be problematic (Mirza-Babaei *et al*, 2011) because they further add to the issues surrounding the subjective nature of different user experiences. Players having different physiological reactions to a game is an inherent part of this type of testing, especially due to the influence that indivdual bodily health can have on results (Biag *et al*, 2019). It is therefore sensible to use a game in which players engage with simple, clearly defined mechanics, to try and make sure that the intrinsic personal bias of participants has a minimal effect on the data produced. Hence a Unity based platforming game will be used for testing because it can be directly intergrated with the system at hand, and the data collected during the sessions can be tied to specific sections of gameplay through triggers.

**Self-report data**

The *In-Game Experience Questionnaire* (iGEQ) is a simple self-report method used for gathering information about a player's experience (IJsselsteijn, 2013). It is a form of feedback that has been tested various times within experiments that involve psychophysiological measures (Nacke, 2009). There are many other ways that non-physiological data can be collected in GUR, for instance through observation or qualitative interviews (Mirza-Babaei, P, *et al*, 2011). These are regarded as some of the most valuable techniques, but they can often provide results that are open to bias, which is already an inherent issue with the biometric data. They would yield data that would be difficult to convert and incorporate into a system like this, furthermore the process would be timely. The iGEQ will therefore be used, which contains fourteen items, all rated using a Likert-type scale of zero to four. These items are paired off between seven components, each of which relate to psychophysiological states.

**Scoring guidelines GEQ In-Game version**

The In-game Module consists of seven components, identical to the core Module. items are used for every component. The items for each are listed below.

Component scores are computed as the average value of its items.

**Competence**: Items 2 and 9.

**Sensory and Imaginative Immersion**: Items 1 and 4.

**Flow**: Items 5 and 10.

**Tension**: Items 6 and 8.

**Challenge**: Items 12 and 13.

**Negative affect**: Items 3 and 7.

**Positive affect**: Items 11 and 14.

**Figure 1** Scoring guidelines for iGEQ (IJsselsteijn, 2013)

Please indicate how you felt while playing the game for each of the items, on the following scale:

| not at all | slightly | moderately | fairly | extremely |
|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 |
| < > | < > | < > | < > | < > |

| | | |
|---|---|---|
| 1 | I was interested in the game's story | GEQ Core – 3 |
| 2 | I felt successful | GEQ Core – 17 |
| 3 | I felt bored | GEQ Core – 16 |
| 4 | I found it impressive | GEQ Core – 27 |
| 5 | I forgot everything around me | GEQ Core – 13 |
| 6 | I felt frustrated | GEQ Core – 29 |
| 7 | I found it tiresome | GEQ Core – 9 |
| 8 | I felt irritable | GEQ Core – 24 |
| 9 | I felt skilful | GEQ Core – 2 |
| 10 | I felt completely absorbed | GEQ Core – 5 |
| 11 | I felt content | GEQ Core – 1 |
| 12 | I felt challenged | GEQ Core – 26 |
| 13 | I had to put a lot of effort into it | GEQ Core – 33 |
| 14 | I felt good | GEQ Core – 14 |

**Figure 2** Likert scale used in iGEQ module (IJsselsteijn, 2013)

During testing, the game will be paused, and these questionnaires will be handed to the player to fill out in relation to the section of gameplay they've just experienced. This can then be cross referenced with the biometric data collected during that section and a correlation between the three sources can be made.
One primary problem with this form of testing is that whilst it informs upon the players experience at specific moments of play, it can be intrusive to interrupt the flow of engagement (*Baig et al*, 2019). This might potentially influence the participants mindset whilst playing but it is favourable over collecting data at the end, when it is difficult for players to pick apart the whole experience.

**Collecting and analysing HR/GSR data**

The cardiovascular system can be measured using a variety of options, many of which can give insight into a player's physiological state. HR data is collected through the use of a sensor that detects heartbeats and then emits signals at an equivalent rate, which can then be used to determine a person's heart rate over time. A focus for this project will be looking at Heart Rate Variability (HRV), which has mainly been correlated with player arousal or boredom (Drachen,A, *et al* 2010). This is used to detect specific changes in the time between successive heart beats, which will be ideal for analysing sections of gameplay. EKG is a more accurate way of gathering this data, through sensors attached to a person's chest, but can be more intrusive to set up for participants (Nacke, 2015) and involves much more expensive equipment.

GSR uses small, finger mounted sensors to measure the skin conductance of players (Mirza-Babaei, P, *et al*, 2011). The production of sweat in the eccrine glands acts as a variable resistor for the sensors, so as sweat rises in a gland, the resistance decreases. This data can then be used to determine arousal in participants (Singh, 2013), in a way that is low cost and non-intrusive.

Both of these forms of data collection will be used as indicators, not for emotion, but for logging phasic based 'micro events' (Soares, R, 2016) on a per individual basis. This will allow for a rational correlation to be drawn between the iGEQ data and clear moments of player arousal (or lack thereof). These events will be logged for each section of gameplay and cross referenced with the inputs from the questionnaire to produce results for each participant. A consistent pattern will then ideally be present between the results of each participant.

**Common issues with testing and gathering accurate data**

There are logistical issues with using finger sensors for GSR (Bergstrom, J *et al* 2014), due to the need for the participants to use controllers. To reduce any distortion from movement, the sensors will be attached to the fingers that grip the lower part of the controller. There are also problems with delays in the data collected, in relation to physiological reactions that a player might have. The change in skin conductivity will usually occur 5 seconds after an event that might trigger such changes (Nacke, L, 2015), and for heart rate it is even longer. This will have to be accounted for in the design of the system.

**How the system will process the raw data and visually present the results**

Once the raw data is collected for each section of gameplay and the questionnaire results are entered, the system will have to collate this data and present results that can be interpreted for correlation. A study by Drachen *et al* (2010) used Person's correlation coefficient to normalise the data sets and display the significance in any correlation between dimensions. HR was recorded as beats per minute and GSR was measured in μS. The results indicated covariance between the data sources and showed signs of correlation for negative affect and tension. This form of data processing cannot be used for this project because it directly calculates results using the raw data of all participants, rather than producing individual results for each participant and then comparing them.

| Physiological measures | Competence | Immersion | Flow | Tension | Challenge | Negative affect | Positive affect |
|---|---|---|---|---|---|---|---|
| HR | -0.36 | -0.43 | -0.25 | 0.37 | -0.31 | 0.24 | -0.42 |
| EDA | -0.08 | -0.23 | -0.24 | 0.02 | -0.18 | 0.38 | -0.20 |

**Figure 3** Pearson's correlation coefficient between physiological measures and GEQ dimensions (Drachen,A, 2010)

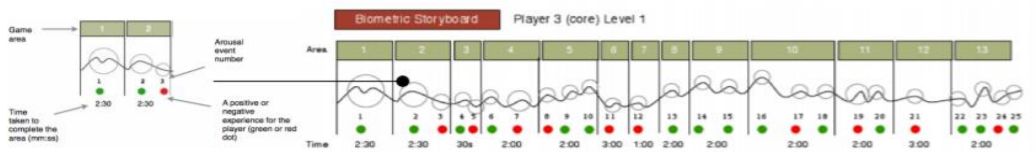


**Figure 4** Biometric Storyboard (McAllister, G. *et al*, 2011)

Instead, the system will detect any significant variations in HR or drops in skin conductivity and log them as micro events. For both of these, this will involve taking a few minutes at the beginning of the sessions to get a base reading from each individual, that will be recorded by the system and used as a normal point for the player's physiological state. The raw biometric data and the micro events will then be presented along with the results of the questionnaire, through a slightly altered version of a Biometric Storyboard (McAllister, G*. et al*, 2011). The storyboards collate biometric data with player reported data in a format that can be easily interpreted.

In this project, the sections of gameplay, along with the time it takes the player to complete them, will be presented in a similar way to McAllister's version. The main difference will be that the graphical data for GSR and HR will be both be plotted on the storyboard, so that any correlation between the two can be easily identified. The micro events will also be highlighted in a similar way to McAllister's version. The iGEQ data will be categorised into the seven components, which will then be assigned either a negative or positive attribute based on IJsselsteijn's method (IJsselsteijn, 2013). These attributes (green for positive and red for negative) will be displayed at the bottom of each section. Once a storyboard is created for each individual, they will all be compared to look for any consistent patterns in the data.

**4. Conclusion and recommendations**

The goal of the project still remains the same, to replicate GUR methodology that cross references different data sources, but the scope will be more refined. The use of biometric storyboards will allow for visual comparison of how consistent the results are when collecting this form of data (the number of participants will therefore have to be large so that this can be done effectively).

The focus will be primarily directed towards using arousal as an indicator of engagement and cross referencing it with direct player feedback, to try and get a consistent pattern in data. Trying to infer specific things about the quality of the game being tested will be avoided.

**5. References**

Baig, M., Kavakli, M. (2019) A Survey on Psycho-Physiological Analysis & Measurement Methods in Multimodal Systems. [online] *Multimodal Technologies and Interaction*. 3(37) pp 1-20 Available from: https://www.mdpi.com/2414-4088/3/2/37 [Accessed 12 November 2019]

Bergstrom, J. (2014) Physiological Response Measurements. [online] In Bergstrom, J., Duda, S., Hawkins, D., McGill, M. *Eye Tracking in User Experience Design.* pp 81-108 Available from: https://www.sciencedirect.com/topics/computer-science/galvanic-skin-response [Accessed 16 November 2019]

Drachen,A., Nacke, L., Yannakakis, G., Pedersen, A. (2010) Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. [online] *Sandbox '10 Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games*. pp 49-54 Available from: http://www.hci.usask.ca/uploads/176-DrachenNackeetal_Correlations.pdf [Accessed 13 November 2019]

Ekman, P., Levenson, R.w., Freisen, W.V. Automatic nervous system activity distinguishes among emotion. *Science. 10*, pp. 1208 (1983)

IJsselsteijn, W. A., de Kort, Y. A. W., & Poels, K. (2013). The Game Experience Questionnaire. [online] Eindhoven: Technische Universiteit Eindhoven. https://pure.tue.nl/ws/portalfiles/portal/21666907/Game_Experience_Questionnaire_English.pdf [Accessed 10 November 2019]

McAllister, G., Mirza-Babaei, P. (2011) Biometric Storyboards: visualising meaningful gameplay events. [online] *In Brain and Body Interface CHI 2011 Workshop.* pp 1-4 Available from: http://physiologicalcomputing.net/bbichi2011/Biometric%20Storyboards%20-%20visualising%20meaningful%20gameplay%20events.pdf [Accessed 17 November 2019]

Mirza-Babaei, P. (2011) *Understanding the Contribution of Biometrics to Games User Research.* [online] In: Think Design Play: The fifth international conference of the Digital Research Association (DIGRA). pp 1-13 Available from:https://www.researchgate.net/publication/289761538_Understanding_the_contribution_of_biometrics_to_games_user_research [Accessed 5 November 2019]

Mirza-Babaei, P., Nacke, L., Gregory, J., Collins, N., Fitzpatrick, G. (2013) How Does It Play Better? Exploring User Testing and Biometric Storyboards in Games User Research. [online] In *CHI 2013: Changing Perspectives, Paris.* pp 1499- 1507 Available from: https://publik.tuwien.ac.at/files/PubDat_222770.pdf [Accessed 17 November 2019]

Nacke, L. Games User Research and Physiological Game Evaluation. (2015) In Drachen, A., Mirza-Babaei, P., Nacke, L. (2018) *Games User Research*. Oxford University Press. pp 63-86

Singh, Navjot. (2013) *Arousal Level Determination in Video Game Playing Using Galvanic Skin Response.* [online] MA, Department of Electrical and Instrumentation Engineering, Thapar University. pp 1- 62 Available from: https://pdfs.semanticscholar.org/d65c/af1945183f379140a2680b6ee8d34d746ffb.pdf [Accessed 17 November 2019]

Soares, R., Siqueria, E., Miura, M., Silva, T., Castanho, C. (2016) Biofeedback Sensors in Game Telemetry Research. [online] In *SBC Proceedings of SBGames.* pp 81-88 Available from:http://www.sbgames.org/sbgames2016/downloads/anais/157482.pdf [Accessed 14 November 2019]

**6. Bibliography**

Certicky, M., Sincak, P., Magyar, G., Vascak, Jan., Cavallo, F. (2019) Psychological Indicators for Modelling User Experience in Interactive Digital Entertainment. [online] *Sensors (Basel)*. 19(5), pp 1 Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6427444/ [Accessed 13 November 2019]

Drachen,A., Nacke, L., Yannakakis, G., Pedersen, A. (2009) Psychophysiological Correlations with Gameplay Experience Dimensions. [online] *Sandbox '10 Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games*. pp 1-4 Available from: https://arxiv.org/ftp/arxiv/papers/1004/1004.0243.pdf [Accessed 11 November 2019]

Sarter, F., Papini, G., Cox, L., Cleland, J. (2018) Methodological Shortcomings of Wrist-Worn Heart Rate Monitors Validations. [online] *Journal of Medical Internet Research*. 20(7), pp 1 Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6048383/ [Accessed 13 November 2019]

**Appendix A: Project Log**

| Pearce Jennings 16000680 | Analysing the use of Biometric Technology in Games User Research | | |
|---|---|---|---|
| Date of the week | Tasks set for the week (bullet points) | Brief summary of outcomes achieved, research or practical aspect completed | Questions arising and/or tasks to be taken forward |
| 23/09/2019 | -Finalise base concept | The original concept was refined | What exact form will the project take |
| 30/09/2019 | -Work on proposal | The proposal was developed further | Narrow the scope for this kind of project |
| 07/10/2019 | -Improve proposal based on feedback | Feedback was implemented | Ensure proposal is at its best quality ready for hand in |
| 14/10/2019 | -Look at what tech needs to be ordered | Arduino and sensors ordered | How will the data be collected |
| 21/10/2019 | -Start research around GUR | Research around UX was done | How does this research influence the direction of the project |
| 29/10/2019 | -Further research around GUR | Limitations with the UX testing process were identified | How will these limitations be avoided |
| 04/11/2019 | -Research report started | Research was undertaken | What are the main area's that the report will cover |
| 11/11/2019 | -Start looking at how the system will be made | Started with working the Arduino IDE and collecting data | How will this be integrated with Unity |
| 18/11/2019 | -Very early stage demo started | Investigation was done into how the Unity libraries will read the data set over from the Arduino IDE | Get solid report draft done |
| 25/11/2019 | -Further research was done | Got a draft for research report | Focus on improving report |
| 02/12/2019 | -Improve report from feedback | Report was finalised | Work further on demo |