GAUGING VARIOUS LEVELS OF VISUAL STIMULUS IN VIRTUAL REALITY TO BETTER ASSIST WITH THE DESIGN PROCESS

Project Ocelot

E.T MARTIN
University of New South Wales, Sydney, Australia
edward.martin@student.unsw.edu.au

Abstract.

This research explored the implementation of biometric monitoring hardware to identify an individual's physiological response to a virtual reality(VR) architectural environment. The participants experienced one architectural setting in VR, four different times, with each time presenting the space in a drastically different rendering style (rendered all geometry realistically, rendered all geometry in white, rendered all geometry in colours, rendered all geometry with textures but the scene lighting is flat and unlit). Using Electrodermal(EDA) Electrocardiogram (ECG) sensors as a method to pinpoint emotional arousal within an individual, the tests concluded that a user's emotional response can differ greatly depending on how a scene is presented to them. From this conclusions can be drawn about which visual style is the most optimal to exhibit unbuilt architecture and which visual style may be detrimental when exhibiting architecture.

Keywords. Virtual reality: Physiology; Architecture; Visualisation; Realism.

1. Introduction: Research Aims and Motivations

Driven by the mass adoption of virtual reality (VR) in the architectural industry, the motivation for this research task seeds from the fact that many architectural practices are currently incorporating VR technology into the design process. Due to VR's ability to offer real life sense of depth and perception, users are able to form a deeper understanding of unbuilt architecture that cannot be translated through traditional methods of visual communications such as floor plans, elevations, and static image computer renderings. The integration of VR into a design workflow also gives the benefit of allowing the architect and the client to experience a space prior to investing money and time into its construction. This allows alterations to be made on a project based on a spatial understanding during the design process.

Due to the onslaught of various VR software and hardware on the market however, there is currently no one singular user experience. With the multitude of ways to render and visualise a scene (photorealistic, un-textured, unlit, etc.), the style in which a space is presented to an individual could affect their overall response to it. It is crucial to identify these discrepancies as VR is often implemented to give an individual their first experience of a yet unbuilt space. If an individual's initial reaction to a virtual environment (VE) can vary greatly depending on how it's visualised, identifying what scenes leave a negative or a positive impact on the user is important as it could affect how receptive they are to the space being presented. This means feedback could suffer as a users opinion on the space is influenced by presentation rather than the physicality. One then needs to categorize how to measure user experience.

With the ability to emulate a real world sense of scale and perception it would seem logical to assess VR through the same methods that built architecture is critiqued. To give context, once a real world project is complete a Post-Occupancy Evaluation (POE) or Design Quality Method (DQM) is commenced. These two processes act as a way to measure the design quality, building performance, and operational efficiency of a building. These tests, while proven to be efficient, have shortcomings due to the nature of surveys as a design critique method. Aspects such as a building's aesthetic design decisions may not be critiqued accurately through a survey due to various sociological reasons. For example, social desirability bias is trait that affects "the tendency of research subjects to give socially

desirable responses instead of choosing responses that are reflective of their true feelings." (Kamakura, W.A., 2010)

In an effort to minimise the effect of these social biases, this research project incorporates biometric monitoring sensors into the feedback process. Implemented correctly, physiological monitors are able to indicate internal changes that the human body goes through in reaction to stimulus. Depending on what physiological attribute is being measured, fluctuations recorded by the sensors can be indicative of how a person is responding emotionally to stimulus. Since biometric sensors are measuring the subject's inherent physiological reaction, the feedback process does not have to directly rely on the subjects verbal feedback that may or may not be biased.

2. Research Observations and Objectives

The object of this thesis task is to analyse individual's physiological reaction to a VR environment to get a deeper understanding about the effects of visual stimuli of the user. The focus on visual stimuli (how the scene is rendered i;e realistic textures and lighting vs unlit white room) stems from partnering with BVN architectural firm, and having got first hand experience on how VR is used effectively within the architectural industry. there is a multitude of programs and software that each fulfill a certain role depending on the type of experience one needs to create. As an example, when using VR strictly as a quick visualisation tool, irisVR for Revit is able to quickly convert your 3D model into a VE giving the user the ability to teleport around a basic rendered environment giving the user a general idea of how the space feels. However custom tailored VR experiences or creating extremely realistic scenes may require the user of much more technical software such as video game engines Unreal Engine 4 and Unity

Due to all these variables and design choices that occurs when creating a VR environment. The end objective is to able to discern how the visual stimuli of scene, whether it be photorealistic, colour rendered, unlit, etc, effects greatly how the user responds to the scene.

3. Research Questions

There are particular elements of this thesis task that stem from a past research experiment titled “DESIGNING BETTER SPACES FOR PEOPLE”(Dias 2014). Dias’ experiment concluded that by physiologically monitoring participants in VR certain architectural elements, such as corridors, open spaces, stairs and high walkways, are able to trigger emotions that indicate whether a user feels positive or negative about a space. In Dias’s experiment all spaces were rendered with flat lighting in neutral grey colours. This was done to minimise the effect of visual stimuli. For this thesis however a crucial consideration lies in understanding the user's response to varying levels of visual stimuli.
Our hypothesis is built on this research that architectural elements in VR have been proved to trigger an emotional response in participants. But if an individual is presented with the same space four different times, with each spaces visual language different, will the individual’s response be different for each example? Is the strength of the participants reaction a good indicator for which scene offers the most engagement, be it negative or positive ?

4. Methodology

The research experiment was developed to explore and measure participants reaction to four different VR architectural spaces all consisting of different levels visual stimuli. It implements knowledge and results based on the findings of past VR physiological experiments such as Dias 2014. It also takes into account a variety of ways a scene can rendered and displayed to inform the final visual styles of the scenes (AEC Magazine 2017). The experiment consisted of three key elements.

1) Constructing a VR environment for participants to experience.
2) Monitoring participants with the aforementioned EDA and ECG sensors.
3) Give participants a written survey to fill out to compare the physiological data with.

5. Background Research

5.1. Literary review

The purpose of this literary review is to establish a concise background for the motivation of this project and validation for key aspects pertaining to it. The basic objective of this research project is to measure how an individual's physiological reactions vary within a architectural VR environment, specifically focusing on how a users response to a space differs depending on the way it is visually presented to them. To begin, it was crucial to establish whether current methods of gauging user experience within the built environment were reliable or not. Anton J. Nederhof's research paper discusses this topic of disingenuous feedback referred to as "social desirability bias". Social desirability bias can be described as letting social psychological pressures influence one's opinion clouding a genuine response. In his paper, Nederhof seeks to offer ways of overcoming this bias through numerous questioning methods. However, individually these methods were deemed ineffective(Nederhof, A.J., 1985). To counteract the discrepancies encountered with post occupancy evaluation surveys led to the exploration of physiological monitoring for this project.

Understanding how biometric sensors work and what the physiological data means was crucial to implementing this technology correctly. To be utilised in the experiment the sensors had to give us biological data that was indicative of what a person was feeling in response to the VE. two types of biometric sensors seemed invaluable for gauging an individual's inherent physical reaction is, Electrocardiography(ECG) and Electrodermal(EDA) monitors

Ther ECG uses three electrodes placed on the skin and records electrical activity generated by heart muscle depolarizations, which propagate in pulsating electrical waves out through the skin. From these electrical signals, heart rate (HR) and heart rate variability (HRV) are able to be measured and recorded. EDA works by attaching two electrodes to the skin and measuring the electrical resistance between them. This resistance decreases or increases depending on how much the user perspires. These reactions are significant as research shows that "Pleasantness of stimuli can increase peak heart rate response. HRV were shown to be a useful feature in emotion assessment" and "Electrical resistance decreases due to an increase of perspiration, which usually occurs when one is experiencing emotions such as stress or

surprise." (Koelstra, S., Muhl, C., Soleymani, M., Lee, J.S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A. and Patras, I., 2012.).

With the benefits of physiological monitoring apparent, the next step was applying this technology to the practical side of the project, the virtual reality environments. Upon further research investigation, numerous studies found were comparable to the objectives of this project. "The Responses of People to Virtual Humans in an Immersive Virtual Environment" was a virtual reality experiment that physiologically monitored participants as they reacted to virtual humans in a virtual environment. The project compared the quantitative data of the biometric sensors against the written responses they gave in a survey. This research project acts as a solid precedent in identify numerous ways of collecting human feedback in a virtual space. Key issues that surfaced included discrepancies between the physiological data and the written responses. This highlights how there is unreliability in achieving coherent feedback, but the addition of more technologically advanced monitoring tools helps to shorten this the gap.(Garau, M., Slater, M., Pertaub, D.P. and Razzaque, S., 2005).

Other previous research that has explored the effectiveness of VR coupled with physiological sensors is a research paper titled "DESIGNING BETTER SPACES FOR PEOPLE"(Dias 2014). This project consisted of running various tests subjecting users to various architectural elements in VR (ramps, stairs, corridors). The physiological attributes being measured in this experiment included ECG and EDA activity. By implementing these sensors participants emotional response was recorded in real time in reaction to specific aspects of the virtual environment and researchers could then deduce whether the architectural elements were eliciting a positive response or a negative one.

5.2. biometric equipment.

To be able to retrieve the physiological data from the participants the employment of a biometric retrieval hardware was required. After extensive research the BITalino (r)evolution Board Kit was found to be the most optimal solution. Designed specifically for consumer use and research based tests (BITalino 2017), the board comes with numerous biometric sensors, specifically ECG and EDA sensors, which are crucial for our experiments. Bitalino also includes free software that works in conjunction with the board to record and save the physiological data.

6. Development of the case study

6.1. Constructing the Virtual environment

To achieve the various visual styles needed for the experiment, the VE had to be created and rendered using a software that could output high quality photorealistic visuals, and offer effective means to curate a VR experience. For these reasons Unreal Engine 4 was used as compared to other programs on the market it boasts high quality realistic rendering options and also integrated VR support so scenes could be prototyped quickly and effectively. The VE consist of a 3D model designed to mimic an interior architectural setting. The model was created as a linear experience so users experience each architectural aspect of the space The stages of the experience are as follows:

1) The participant begins the experiment in a small hallway. (hallway)
2) The participant is then teleported to the end of the hallway into a large open space room with various furnishings and large windows so the room can be lit with realistic sunlight. (door)
3) Participant is teleported to the other side of the room to experience a different vantage point.(Room
4) Participant is teleported to a mezzanine level of the space amongst railings and stairs. (Mez)
5) Finally the participant is teleported to a balcony high up in the space on the precipice of the edge. (Balcony)

figure 6.2. 5 vantage points that lead the user through the experience.

The spaces were curated specifically to encourage an emotional reaction of the user, catering to specific emotional reactions that are elicited when one experiences things such as height, claustrophobic spaces and larger open spaces. Furniture and detailing was added to the room to make the space feel more immersive so that the participants can reach a sense of “being there”. Participants exhibited the space in VR through the oculus rift virtual reality

headset. the oculus rift consists of a Pentile OLED display, 1080×1200 resolution per eye, a 90 Hz refresh rate, and 110° field of view.

6.2. Creating the visual style for each scene

Developing the visual styles was an integral part of the experiment as each one needed to explore a unique visual language that had a reason for being exhibited. The numerous visual styles can be categorised as realistic, indicative, suggestive and abrasive.

The indicative scene exhibits the lowest form of detail. Consisting of no textures or shadows, the scene is lit to mimic natural lighting while all the architectural elements are rendered in a muted white. The purpose for this scene is to act as a simple exhibition of space with no overly distracting visual stimuli.

The suggestive scene uses the same lighting conditions as the indicative scene. However, colour is added to illustrate the materiality of the architectural elements. For example, the floor is rendered in brown to indicate it's a wooden material while the walls are rendered in orange to suggest brick. The scene acted as a stylistic way to exhibit built form.

The realistic scene was developed to exhibit the highest form of realism. The scene included realistic natural lighting, shadows, and high resolution photo realistic textures enhanced through the use of normal maps to give the architectural elements textural sense of materiality. Normal maps allowed for high quality shading to be applied to textures i.e. a wooden surface appears to have divots where the grain runs instead of appearing as a flat image.

The abrasive scene consists of no natural lighting; it instead is illuminated just to show basic elements within the space. All the architectural elements have textures of what would be its real world materiality however they have no normal maps and appear as flat surfaces. This scenes was created to assess whether unnatural lighting conditions and textures would negatively impact how participants felt about the space, as it acts as a stark contrast between the more realistically lit and rendered scenes.

figure 6.2. four visual styles from the top left to bottom right, indicative, suggestive, abrasive , and realistic.

6.3. Curating how people experience the space.
ONEdirection

Dictating how people experienced this study was crucial as there are numerous factors that could affect the outcome of the experiment. It was decided early on that the participants will be teleported through the space and not be given the option to roam around freely. This was done for numerous reasons. Firstly, in order to ensure that each participant had the exact same experience. If participants were allowed to wander freely, it would be hard to quantify their reaction to the architectural elements as they would be able to simply control with features they focus on and interact with. This would make comparing experiences between participants immeasurable. For this reason participants were fixed into one spot and could only use their head to look around in the VE. They were then teleported in scene to each vantage point. Secondly, highly sensitive biometric sensor were used on the participants to record their physiological response. To achieve a more consistent and accurate reading we asked the participants to stand still in one spot. If they were to move around freely the sensors could experience interruption resulting in the data being inaccurate.

The next issue involved choreographing how the scene would executed. Factors included; how long one should dwell in a space before they are teleported to next vantage point, which scene should the user start and finish in, and how should the participants transition between scenes. The decision to set how long each participant experienced each element was deduced heavily based on the participant's physiological response. It was important that the user's initial reaction was caught and then left to plateau to give a reliable indication of how one felt overall about the section. For this reason users were given forty-five seconds in each vantage point as we found that people's initial physiological reactions were quite brief. Therefore, letting them dwell for another thirty seconds allowed for a sufficient amount of data to be collected. On the topic of transitioning between scenes, since each participant spent forty five seconds at each of the five vantage points, they had spent nearly four minutes getting familiar with the one scene. Because participants would become accustomed to a specific visual style, transitioning them into the next scene had to be done gradually as a sudden snap to a new visual stimuli could cause physiological spikes in response to the act of transition rather than the actual stimuli of the new room. To combat this, after a participant was in the final stage of a room, users headset's would fade to black where they would see nothing for fifteen seconds, after which the new scene would fade in. We found this method to be less invasive and less jarring.

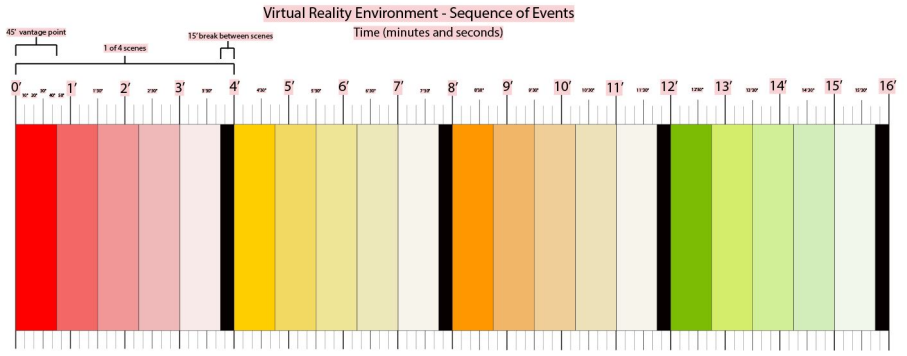

figure 6.3 timeline of how the experiment is run, each colour indicates one scene, gradient shows the progression from vantage points, black is the 15 second interval

The order in which the user experienced the scenes also had to be controlled to minimise the effect of order on physiological responses. For example, it is not unreasonable to propose that a participant's perception of a scene may be affected by their experience of a previous scene. Therefore, if every participant experienced each scene in the same order, the data could be unreliable. To alleviate these discrepancies, participants were split into two separate groups. One group would experience the experiment from realistic, abrasive, suggestive to indicative, essentially going from most detailed scenes to least detailed. The second group experienced the scenes in a random order not based on levels of detail or realism. The order was suggestive, abrasive, indicative, realistic.

6.3. The participants.

The Experiment was conducted on a total of 14 people between the ages of 19 - 25 consisting of 8 males and 6 females. To be eligible to participate in the experiment, it was a requested that all participants be familiar with VR having used it in some capacity before. Since the study was measuring physiological response, it was thought best if participants had previous experience with VR as their reaction should be purely based on what was being presented in VR, rather than based on the act of experience a VR experience for the first time (which is often jarring and exciting for people).

6.4 Experiment results

Out of the 14 participants, physiological data was collected from ten of them (5 males and 5 females) as technical issues with the sensors occurred with the other four.

Analyzing the EDA data showed a multitude of different trends between the participants and the four scenes. To get a better understanding at how strongly subjects EDA peaked for each vantage point in each scene, the data was normalised so percentages could indicate which sections of the scene elicited the largest emotional arousal.

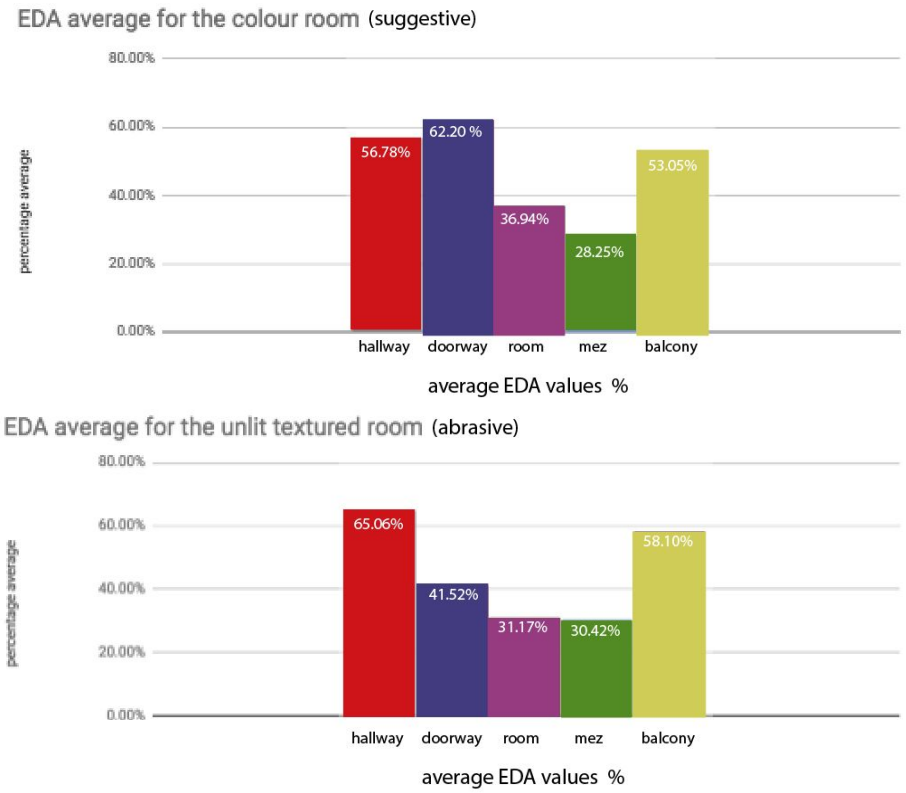

figure 6.4.1 and 6.4.2 . Average EDA values throughout the suggestive and abrasive scenes

By analysing each rooms averages, some deductions could be made based on how people were experiencing the scene. The abrasive scene (unlit textures) and suggestive scene (coloured room) have the strongest peak in EDA at the start of the scene which then starts to dissipate until they reach the balcony where it spikes again to 53.05 %, for the suggestive scene and 58.10% for the abrasive scene(figures 6.4.1 and 6.4.2).

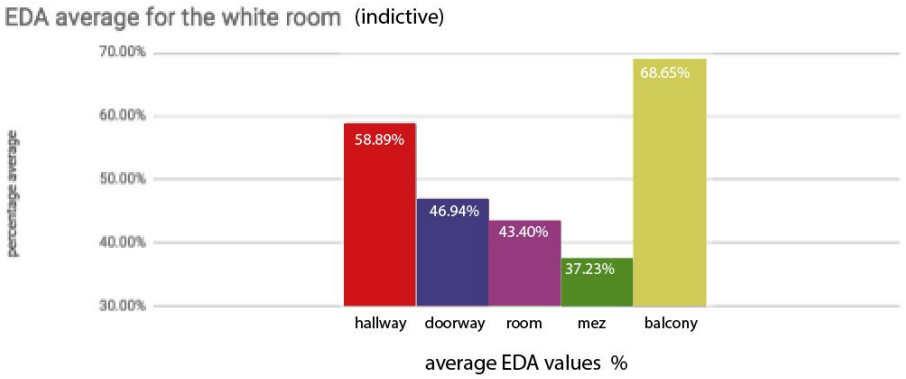

figure 6.4.3 Average EDA values throughout the indicative scenes

The indicative room (rendered all white) showed similar results to the suggestive and abrasives scenes where initial EDA response is strong then dissipates until finally peaking at the balcony with an averaged out value of 68.65%. However, in the indicative scene the balcony ended up eliciting the highest response out of the entire scene.

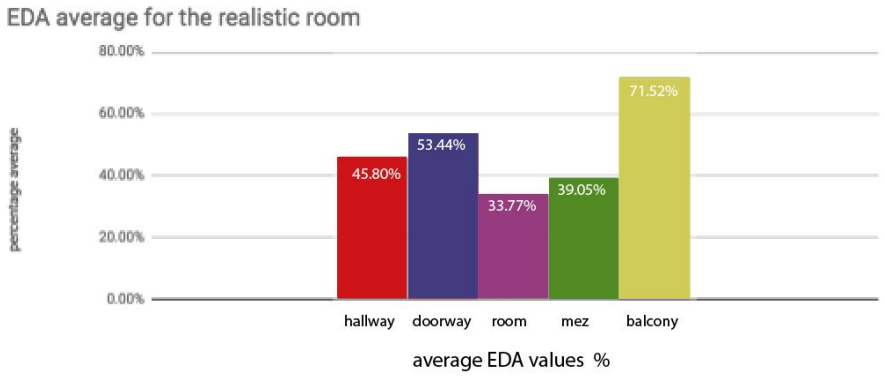

figure 6.4.4 Average EDA values throughout the realistic scenes

The realistic room varied the most out of the four as the average values indicated continuous emotional arousal through the scene and finally peaking at the balcony with a percentage of 71.52%.

These graphs while not indicative of a single person, are helpful in understanding general reactions about the space. What they tell us is that participants a reacting strongly to certain aspects of each scene specifically the first two vantage points and the final one on the balcony. They also help to show which scenes kept the participants continuously engaged. While the

balcony vantage point spikes the EDA response in each scene. The indicative scene and the realistic scene appear to be the most engaging based on how consistently higher the realistic scenes vantage points averaged out and the white room's ability to continue to arouse participants towards the end of the experiment.

To back up the EDA results of the participants and deduce whether emotional arousal was indicative of them enjoying the scene, the participants were given a survey where they were asked to answer 3 questions about each room:

1) Did this scene make you feel curious or apprehensive?
2) Did you feel pleasant or unpleasant in the room?
3) Please rate this room (out of 7)

Participants were also asked to pick their favorite room.

Analysing the survey data, many answers help to back up people's EDA responses to the scenes. About the question pertaining to their favorite room 60% percent they liked the realistically rendered room. This correlates to the deductions that this specific scene holds a higher constant emotional arousal

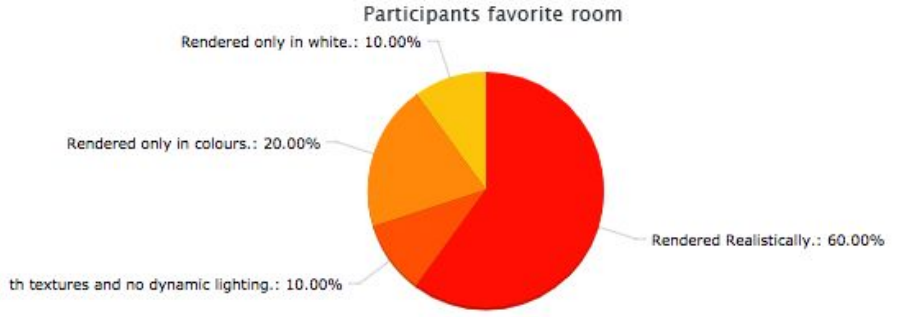

figure 6.4.5 survey results indicating participants favourite rooms

In terms of engagement the realistic room was a favorite among participants with 9 initially saying they were curious about the space, followed by all 10 of them regarding the space as pleasant, the room also scored highest in average level of engagement with the participant with a score of 5.60 out of 7.

Analysing the indicative room, 9 where initially curious about the space, with just over half (6) of the participants finding the space pleasant, the average level of engagement score came to a total of 4.60 out of 7. (figure 6.4.3). Overall this scene seemed to have quite a mild response. Referring to individuals EDA response their trend seem to be emotional arousal going down after the first vantage point, plateauing and then rising up again.

Coupled with the survey data this scene seems to be quite a neutral way of exhibiting form.

The abrasive room was quite a contentious scene with quite a strong opinion conveyed through the survey. 7 out 10 of the participants felt apprehensive entering the scene, a further 7 out 10 answered feeling unpleasant in the scene with a final engagement score of 3.7 out 7. Due to the high level of participants feeling apprehensive within the scene, I compared these survey results with individual participants EDA levels as they went through the abrasive scene to get better indication of their emotional response.

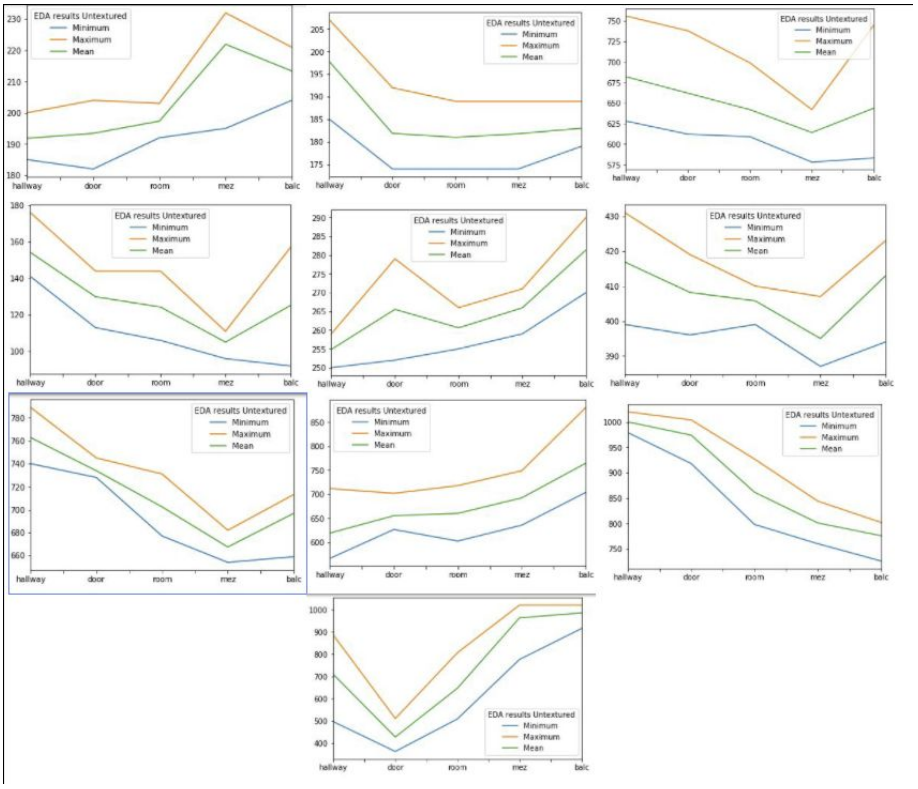

figure 6.4.6 ten participants EDA levels as they went through the abrasive scene

By visualizing each participant's EDA levels (figure 6.4.6) as they proceed through the abrasive scene, 7 participants show a steady drop in emotional arousal as they enter the scene and it most cases continues to drop from one vantage point to next.. This correlates well with the survey data and suggests that participants were not very engaged with the scene and were left apprehensively of the scene

The suggestive scene (rendered in colour) was received quite well. Initially all 10 participants described themselves as curious about the scene, but by the end of the scene a total of 7 described themselves as feeling pleasant about it. Lastly, the scene averaged a score of 4.4 out 7 in response to level of engagement within the space.

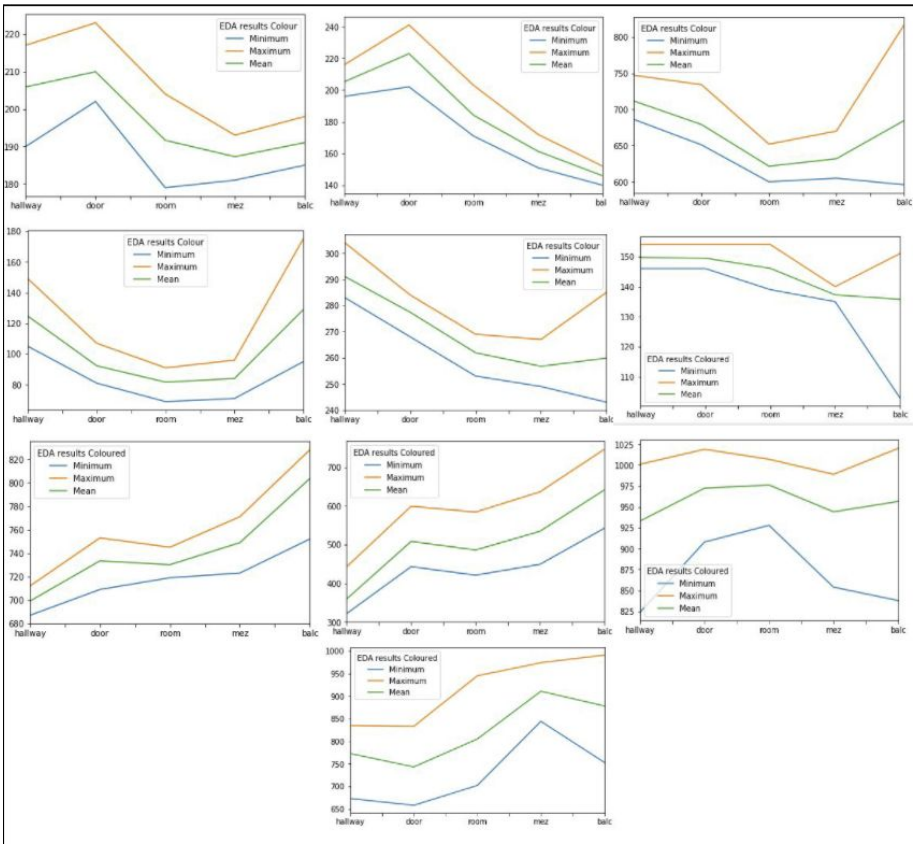

figure 6.4.6 ten participants EDA levels as they went through the suggestive scene

The EDA results indicate the split between participants where half start with a EDA level that suggests they are engaged with the scene that then tapers downwards only to pick up again at the mezzanine vantage point. While the other half see a continuous increase of emotional arousal throughout the scene. Due to the nature of this scene I think the aesthetic nature of cause quite a bit of variability between opinion. The colours are supposed to

indicate materiality but it's possible participants strictly viewed it as a stylistic choice and reacted to that.

7. Significance of Research

This research holds significance as it found immense variability between all the scenes. Each scene consists of the same architectural elements, same scale, and participants progressed through the space at the same rate. However individuals formed completely different opinions on each space while their bodies inherit physiological response differed significantly as well indicating a vastly different attitude towards each scene. This project acts as indicator for variability in understanding architectural elements in VR. When exhibiting a built environment project in VR, architects and designers should strongly consider the style and presentation when immersing a client. Using my project scenes as an example, if I wanted the client to be constantly engaged in the space I would use the most realistic scene as it proved to hold participants emotional response and surveys indicated it was participants favorite room. Another example would be using the indicative room to introduce clients to the space. Emotional arousal for the scene was quite subdued while the majority of participants found the scene pleasant. This scene works well for getting an objective understanding of the space and benefits architects in terms of workload as the model can be exhibited faster due to not having to spend excessive time texturing and building elements in the scene to make it look highly realistic. The abrasive scene acts as the perfect antithesis to these other scenes as most participants had a strong adverse reaction to it. If an individual's first introduction to a space is a negative one, it could lead to a misunderstanding of the space through no fault of the design itself but more of its visual representation.

8. Evaluation of research project

Assessing the project retroactively there are some aspects that could have been altered for the better and some limitations that were experienced throughout the experiment.

The number of participants was quite lacking in terms of what was originally intended. 14 participants with only 10 having usable physiological data is quite a small sample size, higher number would give a lot more varied and interesting data.

More environmental control could be implemented in future research. Tests were carried out on multiple days so temperature and test areas fluctuated between each user test this could of been affecting the users EDA.

EDA as an emotional measurement can be quite volatile as its strictly measuring emotional arousal so the need to clarify specific emotions with the user is needed to pinpoint what they are feeling.

VEs such as the subjective scene were quite hard to get a strong consensus on due to the EDA data being quite decisive between participants. This could be due to the small sample size of the participants or that the room style will only resonate with certain people. In anycase i believe it would be interesting to test another scene with more random less indicative colours to act as contrast between the more subdued colour scheme to see if the EDA responses will be more one sided.

Implementation of the biometrics could have could have been a bit more careful. ECG data was omitted from the results mainly due to how people's movements and cloths kept disrupting the sensors making the data unreliable.

Alterations could be made to the experiment to make it a more immersive experience. Giving the participants freedom to roam around scene untethered instead of ushering them from space to space could be beneficial in understanding emotional response to the space as a whole instead of incremental elements.

9. Conclusion

To conclude this project, numerous questions need to be answered. Does VR have validity in the architectural industry? Yes. From the tests conducted its clear that VR is able to elicit emotional responses within individuals that are able to indicate if users are engaged within the space. Is physiological monitoring a reliable method of gauging a user's reaction to space? somewhat. this research project was able to identify that physiological reaction does correlate to a user's inherit response to a space. Specifying what those response mean however requires a bit of verbal or written communication to understand what's the root of those feelings. Is there a definite way to experience architecture in VR? No. From the tests conducted multiple different visual styles were able to elicit a variety of responses. While some were more consistent than others (realistic scene) i believe further testing need to be done to see how the visual styles compare if the users were to go through the actual built space. The main purpose of this thesis was to identify various response to different VE visual stimulus and in that regard i believe it has succeeded.

References

Kamakura, W.A., 2010. Common methods bias. Wiley International Encyclopedia of Marketing.

Greg Corke. 2017. Virtual Reality for architecture: a beginner's guide. [ONLINE] Available at:http://www.aecmag.com/59-features/1166-virtual-reality-for-architecture-a-beginner-s-guide. [Accessed 31 October 2017].

Nederhof, A.J., 1985. Methods of coping with social desirability bias: A review. European journal of social psychology, 15(3), pp.263-280.

Garau, M., Slater, M., Pertaub, D.P. and Razzaque, S., 2005. The responses of people to virtual humans in an immersive virtual environment. Presence: Teleoperators and Virtual Environments, 14(1), pp.104-116.

Dias, M., Eloy, S., Carreiro, M., Proênça, P., Moural, A., Pedro, T., Freitas, J., Vilar, E., D Alpuim, J. and Azevedo, S., 2014. Designing Better Spaces for People.'. In Rethinking Comprehensive Design: Speculative Counterculture, Proceedings of the 19th International Conference on Computer-Aided Architectural Design Research.

Koelstra, SK, 2015. DEAP: A Database for Emotion Analysis Using Physiological Signals. IEEE transactions of effective computing,, vol 3 part 1, 18-29

bitalino. 2017. bitalino. [ONLINE] Available at: http://bitalino.com/en/.