The Effect of Virtual Environments on Stress Levels

Master Thesis
MTA-191035

Aalborg University
Medialogy

For this project, we have used Unity3D, Autodesk Maya, Gimp and Audacity.

Medialogy
Aalborg University
http://www.aau.dk

AALBORG UNIVERSITY
STUDENT REPORT

Title:
The Effect of Virtual Environments on Stress Levels

Theme:
Stress, Virtual Reality, Environments

Project Period:
Spring Semester 2019

Project Group:
MTA-191035

Participant(s):
Jette Møller Jensen
Michelle Hagemann
Patrick Bang Løyche Lausen

Supervisor(s):
Jens Stokholm Høngaard
Martin Kraus

Copies: 1

Page Numbers: 53

Date of Completion:
May 26, 2019

Abstract:

In Denmark, 250.000-500.000 people suffer from severe stress. Therefore, this report seeks to investigate what affects stress through environmental factors in virtual reality. Two virtual environments were tested against each other; a virtual nature environment and a simple room environment with an implemented Montreal Imaging Stress Task. Three experiments were conducted in relation to this. The results showed no significant difference between the environments. This suggests that environments have a smaller effect on stress levels than could be detected by this experiment. Further research on this subject is however encouraged.

Summary

In this project, it was found that several people suffer from or experience symptoms of severe stress. Therefore, the aim was to find a way to reduce stress.

Through the background research, different ways of measuring stress were looked into, as well as what methods are already used in relation to reducing stress. It was found that nature environments are commonly used as a way to induce relaxation, while also reducing stress. From this, the following problem statement was made:

"What elements in a virtual environment are responsible for inducing relaxation?"

Three experiments were conducted, in which two different virtual environments (a simple room and a forest scene) were tested against each other with an implemented Montreal Imaging Stress Task. For the first experiment, no significant difference were found between the simple room and the forest scene. The null-hypothesis;

"A forest scene is not better at decreasing stress in participants when performing a mental stress task."

could therefore not be refuted. As there was a difference between the baselines and the environments, it could however be found that the MIST did in fact stress people. As the background research showed that nature should have an effect in relation to reducing stress, it was therefore decided to improve upon the environments and repeat the experiment.

For the second experiment, no significant difference between the virtual environments were found. Therefore, the null-hypothesis (the same as for the first experiment) could not be refuted during this experiment either.

For the third and last experiment, audio was implemented in order to create a better immersion within the scenes. The null-hypothesis could not be refuted with the calculated p-values. However, less variables showed a significant difference between the baselines and the environments, compared to the previous experiments. This might indicate that audio could have an effect in relation to reducing stress.

From these three experiments, in which the null-hypotheses could not be refuted, it is not certain that nature has an effect in relation to reducing stress. However, this might have to do with the immersion of the virtual environments. Additionally, there might be indications towards audio being effective in reducing stress, and therefore further research on this topic are to be encouraged.

Contents

Chapter 1

Introduction

In Denmark, in-between 250.000 to 300.000 people suffer from severe stress. In addition, 430.000 people experience symptoms of severe stress everyday. Furthermore, 35.000 people are on sick leave everyday and every fifth person suffering from stress risk losing their job. Lastly, work-related stress costs the Danish society 27 billions every year. Stress is a problem for many people and according to World Health Organisation (WHO), stress will be one of the biggest factors for illnesses by 2020[14].

Stress is defined as a condition in which the individual experiences reluctance and tension[15]. There are two different kinds of stress; short-term stress and long-term stress.

Short-term stress refers to a person going into a state which occurs when the person encounter a burdensome event. In this case, short-term stress helps to deliver better performance. The body releases hormones such as cortisol for the body to use as energy. Once the stressful event is over, the cortisol levels decrease.

The second condition, long-term stress, happens when the cortisol level does not decrease. The symptoms cover behavioural, psychological and physical signs such as insomnia, tiredness and headaches etc. It can affect the immune system as well as the memory, and be the cause for several severe illnesses such as cardiovascular diseases[15].

Since stress is a problem for both health and economics in society, it is important to look into what can be done in order to help reduce this. The following background research questions have therefore been set up:

- How can stress be measured?
- What is the state of the art to induce stress?
- What is the state of the art to induce relaxation?

Chapter 2

Background Research

In this chapter, the different research questions will be investigated; how stress can be measured, as well as the state of art in relation to inducing stress and relaxation responses. These findings will be basis for a problem statement.

2.1 How Can Stress Be Measured?

As mentioned in *chapter 1, Introduction*, it is important to look into what can be done in order to reduce stress. Therefore, it is relevant to research how stress can be measured in order to be able to determine whether a person is stressed or not. Stress has several different indications and symptoms, due to the reactions that happens in the body. When a stressor is introduced, it activates the sympathetic nervous system in the autonomic nervous system. This releases adrenaline and increases cortisol levels, while also increasing heart rate. In order for the body to relax when the threat has been overcome, the parasympathetic nervous system is triggered. The problems arise when there is too much activity in the sympathetic nervous system, compared to that of the parasympathetic nervous system, as this results in adrenaline, cortisol and heart rate not returning to normal levels after stress. Taking these parts of the nervous system into consideration, it means that there are multiple ways to measure stress[11].

One way is to measure the cortisol levels through different bodily fluids such as saliva, urine or blood, as the level of cortisol is increased as part of the activity in the sympathetic nervous system[7]. However, these methods can be rather invasive and require laboratory tests. Another possibility is measuring the Heart Rate Variability (HRV). HRV relates to the variability in heart rate beat-by-beat over time. This method can be used to determine the activity of the sympathetic and parasympathetic nervous systems[7]. This is relevant, since having too much activity in the sympathetic nervous system, without an equal amount of activity in the parasympathetic nervous system, results in stress without a proper relaxation

period[7].

From the measurements given from HRV, statistical time-domain measures, as well as frequency-domain measures, can be calculated. Some of the most common ones are the following:

- Square Root of the Mean Squared Differences of Successive NN Intervals, (r-MSSD)
- Percentage of Consecutive NN Intervals that differ by more than 50 ms (pNN50)
- High Frequency (HF)
- Low Frequency / High Frequency (LF/HF)

If the values r-MSSD, pNN50 and HF increase, it is a sign of relaxation response, while an increase in the LF/HF level shows a stress response for the individual[16].

Other ways to measure stress include qualitative questionnaires that evaluate self-reported data, in order to assess a person's stress levels[17].

2.1.1 Relaxation Time

Having considered the different ways which can be used to measure stress, it is also reasonable to look into how long it takes for the body to relax. This is an important aspect to consider in order to ensure residual stress from other tasks would not carry over into the results during a stress study and thereby affect the reliability and validity of the study.

B. Yu made a study in which they tested upon an auditory interface in relation to relaxation. For relaxation time, they used 3 minutes. They found that a combination of instrumental music and "nature white noise" had potential for being relaxing[19].

Another study made by Soyka et al. made a virtual underwater environment in which the participant had to breathe in rhythm with an animated jellyfish going up and down in the underwater scene. For relaxation time, they used 10 minutes. Their findings showed a similar decrease in stress levels between the environments they were testing[13].

From the found research, it can be concluded that a relaxation time within 3 to 10 minutes should be sufficient to relax people.

2.2 What Is the State of the Art to Induce Stress?

To test the effectiveness of reducing stress, it is important to have a stress response from the participants. Therefore, it is relevant to look into what methods that can be used to induce stress.

One way of inducing stress is the Trier Social Stress Test (TSST)[8]. This test consists of several individual parts. First, the participant will be waiting in one room for 45 minutes. After waiting, the participant is introduced to the test in another room. They are given 10 minutes to prepare a speech as to why they would be an ideal candidate for their ideal job. When the 10 minutes have past, the participant will be told to present their speech in front of a camera for five minutes. If they stop talking during the speech, the participant will be reminded that they still have time left. After the five minutes, they are asked to subtract 13 from 1022 sequentially for five minutes and give their results out loud. If they do not give the correct answer, they are asked to start over from 1022. Then, the test samples will be extracted from the participant. This concludes the test, after which the facilitator will explain the actual test[8]. According to Dedovic et al., this test has been repeatedly successful, inducing 2-4 fold increases in cortisol levels.

Another way to induce stress is the Montreal Imaging Stress Test (MIST)[3] which is a method derived from the previously mentioned TSST. This method is primarily based on the math part of the TSST method by using an algorithm to present different math tasks of random difficulty. Time limits are used to induce the stress by being slightly lower than the average response time of the participant. The participant will first take part in a training session in which the math questions are introduced without a time limit. This is followed by an imaging session to introduce the time limits. The participant will be instructed to have a performance of 80% and be reminded during the test. According to Dedovic et al.[3], this test is found to have an increase in cortisol levels by 50-100%, relative to the baseline[3].

Looking at these two methods used to induce stress, there are advantages and disadvantages with them both.

TSST has as mentioned a very high success rate, repeatedly getting 2-4 fold increases in cortisol levels during the procedure. It is one of the most widely used ways to induce stress[3]. However, the procedure is exceedingly long, taking multiple hours to complete.

MIST is, as mentioned earlier, derived from TSST. It has been shown to induce an increase in cortisol levels by 50-100%, relative to the baseline. Compared to TSST, it is less time-consuming.

2.3 What Is the State of the Art to Induce Relaxation?

Bratman et al.[2] made a study in which they compared the effect of an urban environment against a nature environment. It was found that the natural environment had a bigger effect in decreasing the stress level of the participants[2].

Furthermore, Jiang et al.[6] conducted a study based on the results of Bratman et al. In this study, they tested the optimal amount of nature elements in an environment. They found that a bigger amount of nature elements had an increased

impact on decreasing stress levels.[6]

Other studies have also been done on the effect of different natural environments in regards to stress. Mahalil et al.[9], Soyka et al.[13] and Thoondee et al.[18] made studies in which they tested different simulated nature environments.

Mahalil et al. tested cognitive behavioural therapy methods in which a participant was sent through a guided meditation. The participant had to imagine being in a natural environment. This guided meditation was tested against a virtual environment which simulated a natural environment with the same guided meditation. In this paper, it was found that the virtual environment was better at relaxing people than the cognitive behavioural method[9].

In the study by Soyka et al. in which they made an underwater environment and used breathing exercises to relax, they tested it against a white room with the in which the participant was asked to perform the same task. They found that the two different environments had similar decreases in stress levels, but that the underwater environment was preferred[13].

Lastly, Thoondee et al. made a virtual environment depicting a field in which the participant could walk around and interact with a few objects. The environment was tested through self-reported ratings of stress levels before and after experiencing the virtual environment[18].

Looking at these different ways in which it has been tried to induce relaxation through virtual environments, some aspects should be taken into consideration in regards of reliability and validity.

Most of the researched papers primarily looked into the effect of one environment. However, they do not procure any research or justifications as to why they chose these environments. Only Jiang et al. tried to test what element of an environment was useful for inducing a relaxation response[6]. However, they only tested one element, in this case being the amount of trees in the environment. The others chose natural environments but had little backing as to why these environments should induce relaxation.

It should also be noted that Mahalil et al., Soyka et al. and Thoondee et al. did not induce stress in their participants using any standardised stress tests. This could mean that their results are not valid since the participants might not be stressed during the test[9, 13, 18].

Additionally, Soyka et al. utilised breathing exercises for their test[13] which in themselves might have an effect on decreasing stress[4]. Therefore, their results might not be valid, since these might not relate to the effect of the environments only.

2.4 Problem Statement

In the background research, different research questions were looked into. As mentioned in *section 2.1, How can stress be measured?*, stress can be measured through various methods. The most precise method would be measuring cortisol levels in the individual. However, as this is measured through bodily fluids such as blood, urine or saliva, They will both require laboratory tests and can have the risk of being too invasive for the participants. These methods are therefore outside the scope of this project. Stress can also be measured through HRV that considers different calculated measurements in order to determine the level of stress. Because of the less invasive nature of using HRV measurement, this method is therefore be the optimal measure of stress for this project.

Different ways of inducing stress have been looked into in *section 2.2, What is state of the art to induce stress?*. Although TSST was found to have the higher increase in stress levels, it is also time-consuming, compared to the MIST method which still provides good results in a shorter test duration.

Lastly, there are different ways for inducing relaxation. However, according to the research found in *section 2.3, What is the state of the art to induce relaxation?*, using environments to induce relaxation proved to be the most common method.

While investigating the different research questions, it was not possible to find a project which had looked into what parts of a natural environment are relaxing to people. This leads to the following problem formulation:

"What elements in a virtual environment are responsible for inducing relaxation?"

Chapter 3

Experiment 1

This chapter describes the details and procedures of the first experiment.

3.1 Design & Implementation

In the background research, Bratman et al.[2] found that nature had a decreasing effect on stress levels. However, in this paper, the environments which were tested were real. This means that there potentially could be many confounding variables which are hard to control.

For this experiment, it was therefore chosen to test two environments; a simple room and a forest scene. These two test environments were chosen in order to test whether nature had an effect on the stress levels or not. In order to be able to control every part of the test environments, it was decided to implement them as *Virtual Reality*(VR) experiences. The further details of the designs and implementation for the test environment have been elaborated upon in the accompanying worksheet for this report.

As can be seen in figure 3.1, the simple room includes two pieces of furniture; a chair and a table. Its measurements are 5x5 meters, and the room has no doors or windows. This is due to wanting to make a contrast to the other test environment, the forest scene, that has a wide space.

Figure 3.1: This figure shows the simple room.

The forest scene consists of a wide plane of 1000x1000 meters on which different 3D models have been placed in order to set the scene (see figure 3.2). These models include different variations of trees and a stone. Along with this, the scene has been decorated with a terrain, a grass shader and a wind zone to increase the realism in the scene.

Figure 3.2: This figure shows the test environments including the forest scene.

3.1.1 Stress Task

In order to ensure a stress response in the participants, a VR implementation of the MIST was created. The MIST has five stages in which math questions of increasing difficulty are presented. The math questions of the different stages of difficulty have the following rules[1]:

- **Stage 1** - 2 1-digit integers (0-9), only + or -, solution 0-9 (1 digit)
- **Stage 2** - 3 1 digit integers (0-9), only +,- (no repeated operations, order of operations random)
- **Stage 3** - 3 integers, 1-2 integers are double digits (0-99), +,-,* (no repeated operations, selection and order of operations random), solution 0-9 (1 digit)
- **Stage 4** - 4 integers, 1-2 integers are double digits (0-99), +,-,* (no repeated operations, selection and order of operations random), solution 0-9 (1 digit)
- **Stage 5** - 4 integers, 1-4 integers are double digits (0-99), *, / are a must, then either + or - (order of operations random), solution 0-9 (1 digit)

In order to create this, a User Interface(UI) is needed, with controls for VR. The UI for this implementation can be seen in figure 3.3.

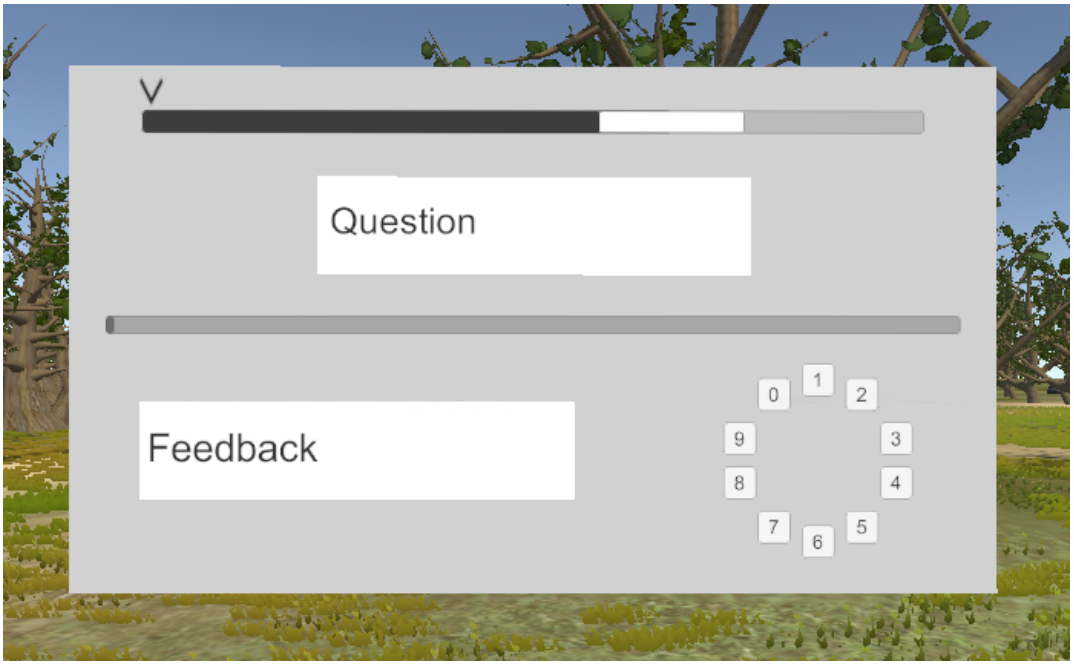

Figure 3.3: This figure shows the UI for the stress task in the VR settings.

On the top of the UI, a performance bar has been implemented. This performance bar shows the average performance for all participants (gathered during the trial run) whereas the 'V' represents the participant's performance in real-time. Performance is here defined as the number of correctly answered questions. This results in the participant usually performing worse than the average, which should be stressful to the participant.

The second element from the top is the text box that shows the current question. This will display a random math problem and the result is displayed as a "?".

Below the presented math question is another bar. This bar indicates how long the participants have to answer the question. When this time has run out, a new question will be displayed.

In the bottom left of the UI is a feedback field which can display "*Correct!*", "*Incorrect!*", or "*Timeout!*".

Lastly, the wheel of numbers in the bottom right allows the user to choose an answer. The number which is currently chosen has a dark outline.

For the implementation in VR, controls were needed as well. Pressing the left or right side of the touch pas of the HTC Vive wand changes the highlighted number counter-clockwise or clockwise, respectively. In order to select a number, the trigger is used.

In order to create math questions, it was required for the implementation to create (or simulate the creation of) random math questions. For the first difficulty, a random number between 0-9 was found (using a function for generating random numbers) and then a random result was also found between 0-9. From that, it is possible to find the result using algebra.

For the second stage, there were 3 integers but still only "+" and "-". This meant that the implementation from the first stage could be reused by just adding another random variable.

The third stage started incorporating multiplication, as well as higher numbers. This was done in order to make it more difficult. By making sure that the result of the multiplied numbers were never more than 99 above or below the randomized result, it was possible to ensure that a solution would always be made which satisfied the equation. If it was impossible to solve the equation with the currently random integers, the function was recursively called again, randomizing new numbers.

Another integer was added in the fourth stage, but by using the same procedure as with the third difficulty, it was still possible to generate solutions that would work (being in-between 0 and 9).

The final stage included division. Although it was possible to generate equations, the complexity which was required made the implementation stop for several seconds. This meant that it was not useful for a real-time implementation. Therefore, the solution that was found, was to generate 70 equations (this amount

seemed sufficient for the final stage which only lasted a minute). These were added as a list to the implementation and a random equation would be chosen every time a new equation was needed. The answers for each equation was compiled in another list, so that it was possible to get both an equation and an answer.

3.2 Evaluation for Experiment 1

3.2.1 Purpose of the Experiment

According to Bratman et al. nature should have an effect on decreasing stress levels. Therefore the main purpose of this test is to see whether proper measurements of stress levels can be made. It would therefore be interesting to investigate whether these findings could be replicated or not. As this would allow further testing into what elements of an environment that affects stress levels. Along with this, MIST is not usually implemented through VR, so it is relevant to test whether this VR implementation is capable of inducing stress or not.

For this experiment, the following null-hypothesis was made:

"A forest scene is not better at decreasing stress in participants when performing a mental stress task."

3.2.2 Participants

For this test, 10 participants were recruited. Of the participants, 9 were male and 1 were female. All participants were recruited at the CREATE campus of Aalborg University.

3.2.3 Setup & Materials

This section includes the setup and materials of the test. For the test, the following equipment was used.

- Two computers
- Camera and tripod
- OBS Studio
- HTC Vive
- Zephyr BioModule heart rate monitor
- Zephyr BioHarness
- Stop watch

For the test one computer was used to run the games and control the HTC Vive. OBS Studios was used on this computer to record the game lay of the participants. The second was used for noting the time from the stop watch at each step of the test, e.g. starting the training session or the start of a break. The camera and tripod was used to record the test. The Zephyr BioHarness was used to secure the Zephyr BioModule heart rate monitor on the participants' chest. The Zephyr BioModule heart rate monitor was used to record the heart rhythm of each participant. The chair was used for letting the participant sit and relax on in the two 5 minute breaks in-between testing the environments.

The setup can be seen in figure 3.4.

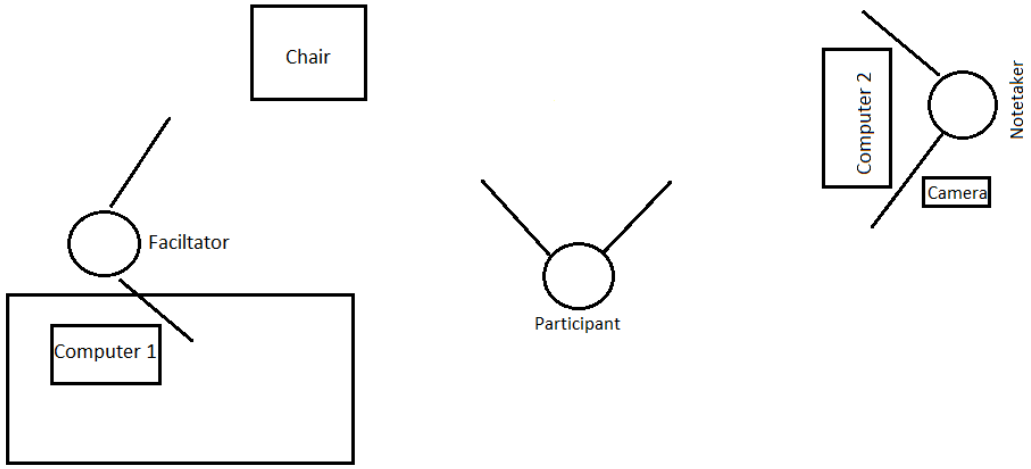

Figure 3.4: This figure shows the setup for when the participant is doing the MIST task in the VR environment.

3.2.4 Procedure

In this section, the procedure of the experiment will be explained in further detail. First, the participant would be greeted and told what the test would contain. They were however not informed that they would be stressed on purpose during the test, since this could affect results negatively, in terms of validity. The participants would then be asked to sign a consent form. If consent was given the camera would be turned on. The participant would be asked to put on a heart-rate monitor and would be told what they would have to do when they entered the game, before they were given the VR *Head-Mounted Display* (HMD) and controllers.

The facilitator would then start the training session from the computer. This training session consisted of the MIST UI presented in a white space. The participant would go through the training for 5 minutes. Afterwards the participant

would be asked to sit down and relax for further 5 minutes. The participant would then be given the VR HMD again and the first of the two environments would be displayed. The facilitator would then explain the difference in the UI to the participant. The order in which the environments were given to the participant were randomized. After the participant had been in the first environment for 5 minutes, they would again be asked to relax for 5 minutes and be put in the second environment. Having had 5 minutes in the second environment, the test would be concluded. The participant would be informed of the purpose of the test and that the intention was to stress them. The participant would be thanked for their help and offered a cookie.

3.2.5 Problems

While conducting the experiment, some issues occurred which meant that some tests had to be redone. This was due to the heart rate monitor not recording the measurements correctly. There can be various reasons as to why this happened, including incorrect placement of the heart rate monitor, clothing obstructing the connection between the skin and the heart rate monitor, tightness of the strap securing the heart rate monitor, the amount of charge on the heart rate monitor, as well as faulty equipment. This meant that the heart rate data for 6 out 10 of the participants were not complete and therefore 6 tests had to be redone with new participants. These tests were conducted after several evaluations of the heart rate monitor had been done. Furthermore, the heart rate data for each participant was examined between each test.

3.3 Results

In this section, the results from the experiment will be presented. As mentioned in *section 2.1*, different variables were looked into, in relation to how to measure stress through HRV; LF/HF, HF, pNN50 and r-MSSD. An one-way repeated measures Analysis of Variance (ANOVA) was used to look at the data for each variable. Baseline was measured between the training session and the first environment, while Baseline2 was measured between the two environments.

According to t-tests, comparing the order in which participants experienced each condition, order did not have any significant effect on the HRV measures of HF and LF/HF. However, for the pNN50 measurements, order showed a significant difference for the training session (Forest first mean = 6.04, Room first mean = 27.1, $p = 0.043$). The r-MSSD values also showed a significant difference in the forest, simple room and training scenario (Forest first mean = 38.04, Room first mean = 25.62, $p = 0.03197$, Forest first mean = 32.88, Room first mean = 25.62, $p = 0.02692$, Forest first mean = 30.42, Room first mean = 45.88, $p = 0.04316$).

Figure 3.5 shows a box plot of the LF/HF data. It includes the different conditions made for the one-way repeated measures ANOVA which was used; AvgBaseline, Baseline, Baseline2, Forest, Room, Training. AvgBaseline is the average of the two baselines. Looking at the medians in the boxplot, the medians for the conditions are overall the same. The ranges vary in size for every condition, Forest being the biggest and Baseline being the smallest. Out of all six conditions, Forest is the only one without an outlier. Lastly, looking at the lower and upper quartiles for the box plot, the data varies the most for Forest, Room and Training.

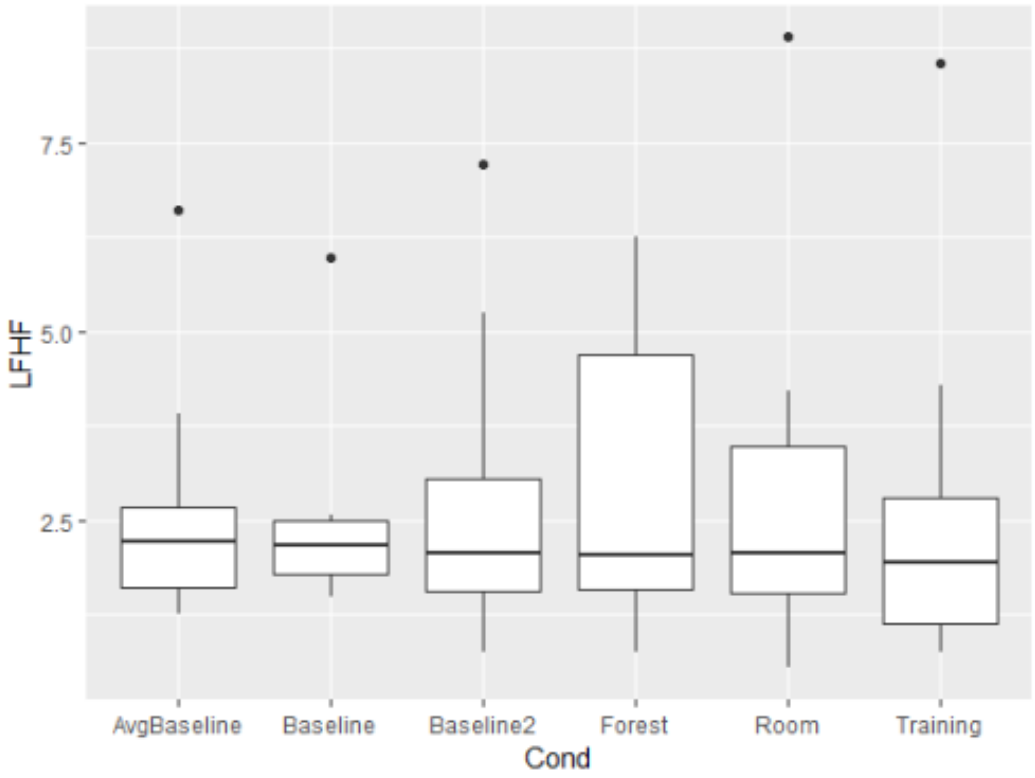

Figure 3.5: In this figure, the boxplot for the LF/HF variable can be seen. The thick horizontal line represents the median of the data, the white boxes represents the upper and lower quartile of the data. The whiskers represents the the biggest and smallest values of the data. The small dots represent the outliers in the data.

Figure 3.6 shows a box plot of the HF data. It includes the different conditions made for the one-way repeated measures ANOVA which was used; AvgBaseline, Baseline, Baseline2, Forest, Room, Training. AvgBaseline is the average of the two baselines. Looking at the medians in the boxplot, the medians for the conditions AvgBaseline, Baseline and Baseline2 are overall higher than the medians for the conditions for the Forest, Room and Training. Among the three mentioned last,

Forest has the lowest median. The ranges vary in size for every condition, Baseline2 being the biggest and Forest being the smallest. Out of all six conditions, Baseline2 and Forest have no outliers whereas the other conditions do. Lastly, looking at the lower and upper quartiles for the box plot, the data overall vary a lot for the three baseline conditions. Forest, Room and Training have smaller variation in their data.

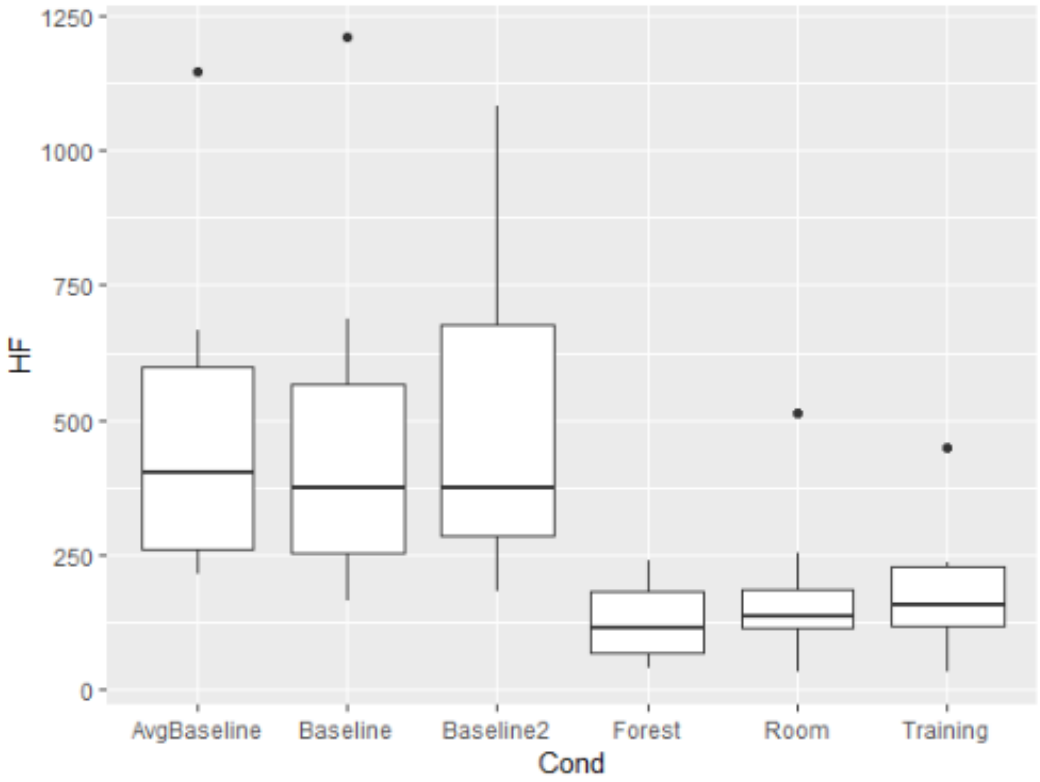

Figure 3.6: In this figure, the boxplot for the HF variable can be seen. The thick horizontal line represents the median of the data, the white boxes represents the upper and lower quartile of the data. The whiskers represents the the biggest and smallest values of the data. The small dots represent the outliers in the data.

Figure 3.7 shows a box plot of the pNN50 data. It includes the different conditions made for the one-way repeated measures ANOVA which was used; AvgBaseline, Baseline, Baseline2, Forest, Room, Training. AvgBaseline is the average of the two baselines. Looking at the medians in the boxplot, the highest median is Baseline. Among the three mentioned last, Forest has the lowest median. The ranges vary in size for every condition, Baseline being the biggest and Forest being the smallest. Out of all six conditions, all baseline conditions have no outliers whereas the other conditions do. Additionally, the three conditions Forest, Room and Training each have two outliers. Lastly, looking at the lower and upper quartiles for the box plot, the data overall vary a lot for the three baseline conditions.

Forest, Room and Training have smaller variation in their data.

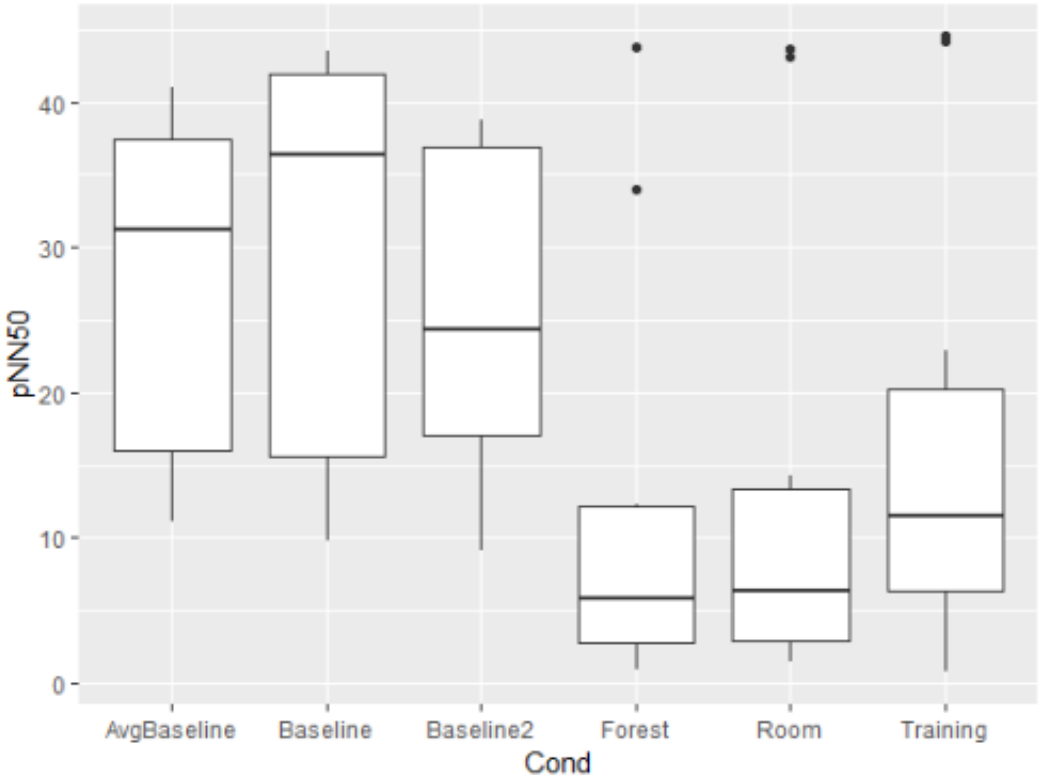

Figure 3.7: In this figure the boxplot for the pNN50 variable can be seen. The thick horizontal line represents the median of the data, the white boxes represents the upper and lower quartile of the data. The whiskers represents the the biggest and smallest values of the data. The small dots represent the outliers in the data.

Figure 3.8 shows a box plot of the r-MSSD data. It includes the different conditions made for the one-way repeated measures ANOVA which was used; AvgBaseline, Baseline, Baseline2, Forest, Room, Training. AvgBaseline is the average of the two baselines. Looking at the medians in the boxplot, the medians for the conditions AvgBaseline, Baseline and Baseline2 are overall higher than the medians for the conditions for the Forest, Room and Training. Among the three mentioned last, Forest has the lowest median. The ranges vary in size for every condition, Baseline being the biggest and Forest being the smallest. Forest is the only condition with outliers. Lastly, looking at the lower and upper quartiles for the box plot, the data overall vary a lot for the three baseline conditions, as well as Room and Training. Forest has a smaller variation in its data.

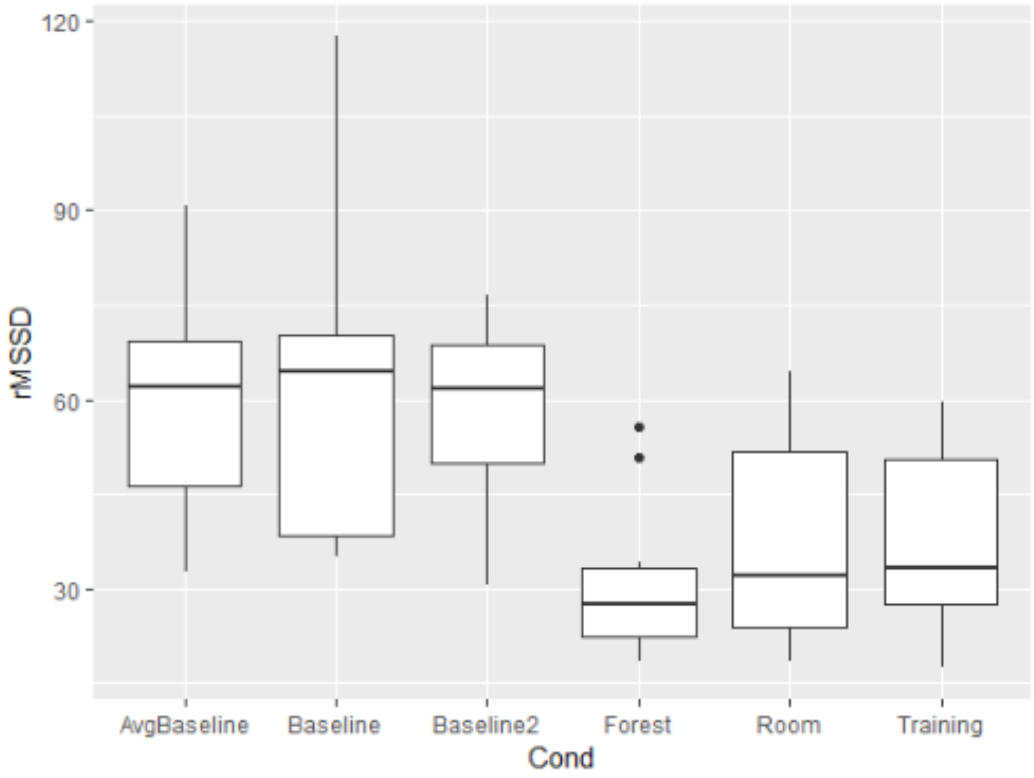

Figure 3.8: In this figure, the boxplot for the r-MSSD variable can be seen. The thick horizontal line represents the median of the data, the white boxes represents the upper and lower quartile of the data. The whiskers represents the the biggest and smallest values of the data. The small dots represent the outliers in the data.

3.3.1 Analysis of Results

In this subsection the presented results will be analysed.

For the LF/HF variable an one-way repeated measures ANOVA showed that the effect of the environment was not significant on LFHF values, $F(5, 54) = 0.10$, $p = 0.99$.

An one-way repeated measures ANOVA showed that the effect of the environment was significant for the HF values, $F(5, 54) = 5.9$, $p = 0.00019$. The following post hoc analysis using the Tukey HSD method showed that the forest environment ($M = 125.28$, $SD = 66.98$) was significantly different from the baseline ($M = 473.69$, $SD = 289.79$), $z(45) = -5.04$, $p < 0.0001$. The simple room environment ($M = 173.94$, $SD = 134.33$) was likewise found to be significantly different from the baseline ($M = 473.69$, $SD = 289.79$), $z(45) = -4.34$, $p = 0.0002$. However, no significant difference were found between the simple room environment ($M = 173.94$, $SD = 134.33$) and the forest environment ($M = 125.28$, $SD = 66.98$), $z(45) = 0.7$, $p = 0.98$.

In regards to the r-MSSD values, an one-way repeated measures ANOVA showed that the effect of the environment was significant, $F(5, 54) = 6.1$, $p = 0.00015$. The post hoc analysis showed that the forest environment ($M = 31.02$, $SD = 12.73$) was significantly different from the baseline ($M = 59.8$, $SD = 17.69$), $z(45) = -4.6$, $p < 0.001$. The simple room environment ($M = 36.84$, $SD = 16.84$) was likewise found to be significantly different from the baseline as well ($M = 59.8$, $SD = 17.69$), $z(45) = -3.67$, $p = 0.0033$. However, as with the HF values, no significant difference were found between the simple room environment ($M = 36.84$, $SD = 16.84$) and the forest environment ($M = 31.02$, $SD = 12.73$), $z(45) = 93$, $p = 0.94$.

Considering the pNN50 values, an one-way repeated measures ANOVA showed that the effect of the environment was significant, $F(5, 54) = 3.0$, $p = 0.018$. This was followed up by a post hoc analysis which showed that the forest environment ($M = 12.28$, $SD = 14.72$) was significantly different from the baseline ($M = 27.77$, $SD = 11.72$), $z(45) = -4.28$, $p < 0.001$. The simple room environment ($M = 13.43$, $SD = 16.3$) was also found to be significantly different from the baseline ($M = 27.77$, $SD = 11.72$) , $z(45) = -3.96$, $p = 0.00103$. However, no significant difference were found between the simple room environment ($M = 13.43$, $SD = 16.3$) and the forest environment ($M = 12.28$, $SD = 14.72$), $z(45) = 0.32$, $p = 0.99953$, in relation to the pNN50 values.

3.4 Partial Discussion

For this experiment, there are various aspects needed to considered, both related to the reliability and validity of the results.

There are some different issues which should be taken into consideration in relation to the procedure of the test. To start off with, the facilitator needed to secure the heart rate monitor on the participant. For the best measurements, the heart rate monitor needed to be attached underneath their clothing. Some people might not be comfortable with having to expose parts of themselves. This might make the participant nervous and therefore affect the results negatively. In terms of the procedure, it was necessary to have a break in-between each of the different tests in order to have a baseline of their heart rate. However, these breaks might not result in a correct baseline of their heart rate, as some people might be uncomfortable while waiting, or feeling uncomfortable with the silence.

It is also necessary to consider the relaxation time in-between the testing the different environments. In the case that the relaxation time is too small, there might be a carry-over between results, as the participants have not had enough time to relax. Lastly, each session was recorded which might also affect the mental state of the participants negatively.

The reliability and validity of the results might also be affected by the participants recruited for the experiment. First of all, some of the people were known to

the facilitators, which might have created a bias in terms of wanting to give good results. Another aspect to consider in terms of the participant is that the knowledge and experience with VR might vary from each participant. This might affect their performance during the tests and therefore affect the results negatively. Lastly, some of the participant mentioned that they felt like they were being stressed on purpose - a suggestion which the facilitators did not give a validation to, in order to avoid affecting the test results. However, feeling that the stress was intended might have their performance and reactions during the tests.

In terms of the equipment used for the test, there might also be different sources of error. Considering the heart rate monitor, it might not have been correctly secured for each session which might cause faulty measurements from the monitor. Some participants also mentioned that the controllers reacted differently from the intended input given by the participant. This affected their performance as they might not be able to answer the math question correctly within the time limit, and therefore might make them more stressed.

Lastly, it should be mentioned that there might not be a sufficient amount of participants for the tests (as mentioned, a total of 10 participants) which might affect the reliability of the results.

3.5 Partial Conclusion

For this experiment, it was chosen to test two environments; a simple room and a forest scene. These two environments were chosen in order to find out whether nature had an effect on the stress levels or not. These environments, as well as the MIST, were implemented in VR for this experiment.

Looking at the analysis of results in *section 3.3.1*, the null-hypothesis could not be refuted, as there was no significant difference in stress levels between the two environments. However, it should be mentioned that the MIST VR implementation did in fact stress the participants, as a significant difference in the stress levels between the baselines and the two environments could be found.

Chapter 4

Experiment 2

In this chapter, the second experiment will be explained in further detail.

4.1 Purpose of the Experiment

As mentioned in *chapter 3, Experiment 1*, the null-hypothesis which had been set up, could not be refuted. It could however be confirmed that the implemented MIST did stress the participants.

The findings from the background research showed that natural scenery should have an effect in regards of decreasing stress[2, 6]. There could potentially be many reasons for this discrepancy in data. From the results of Jiang et al. they found that more nature elements would result in better relaxation. Therefore, it could be hypothesised that the forest environment of Experiment 1 was not successful in simulating nature properly.

For the second experiment, it has therefore been chosen to improve upon the already implemented environments, partially based on comments from the participants, especially for the forest scene. Additionally, it was decided to incorporate the MIST better in each of the environments, so it feels more natural.

As this experiment is based on the same assumptions as Experiment 1, it was decided to use the same null-hypothesis as that experiment:

"A forest scene is not better at decreasing stress in participants when performing a mental stress task."

4.2 Design & Implementation

This section will explain the changes made to the different environments for the experiment; the simple room and the forest scene.

As can be seen in figure 4.1, the MIST implementation has been incorporated into the environment by implementing a TV screen in which it will be shown to the participant. This means that the program in itself has not been changed, but its setup in the scene has been changed.

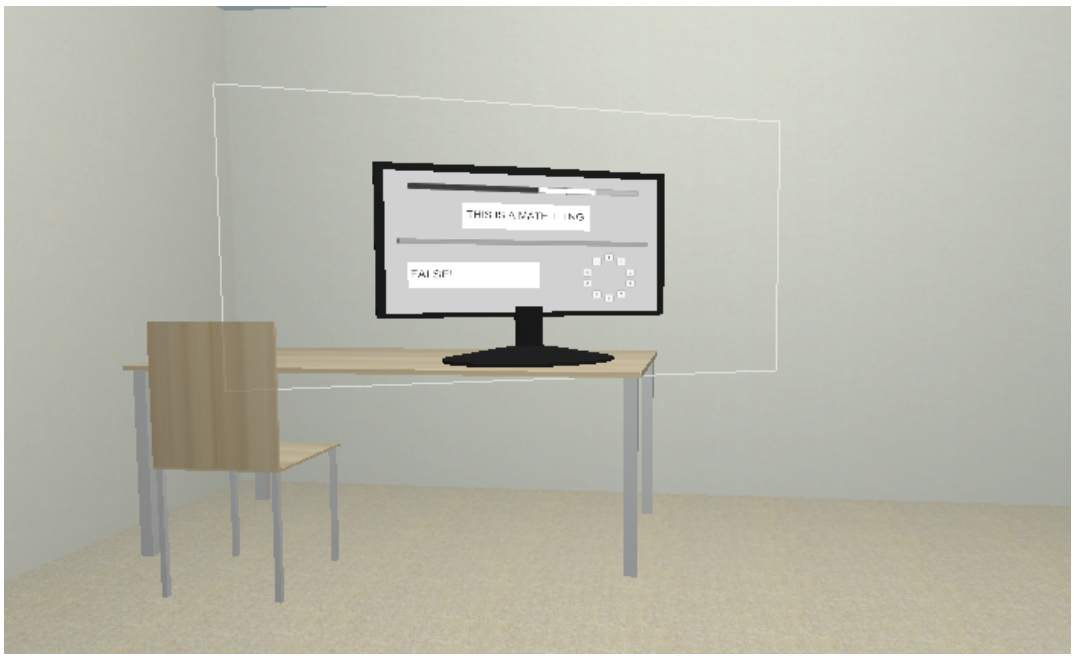

Figure 4.1: This figure shows the second version of the simple room environment with a better incorporation of the MIST.

The MIST implementation has been incorporated more into the forest scene as well, here by putting it up on a sign of wood (see figure 4.2). As with the simple room, the program has not been changed, merely its setup. Additionally, the environment has been improved by various changes. Branch textures replaces the leaf textures on the trees, in order to give them more volume and look more realistic. Some participants mentioned during Experiment 1 that the grass was unsettling to them because it followed the camera (so when the participant moved their head, it would follow as well), and was therefore changed together with the overall structure of terrain (different height levels was added).

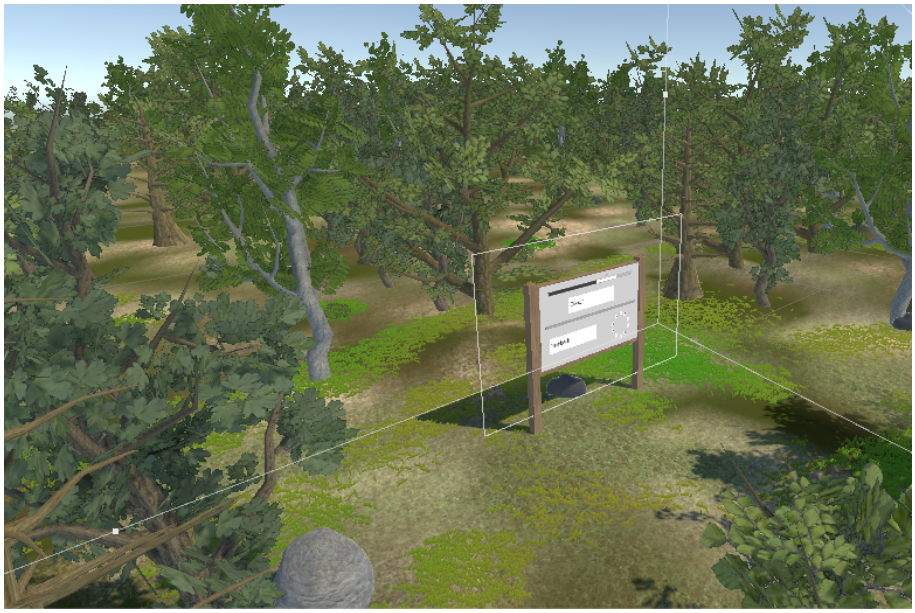

Figure 4.2: This figure shows the second version of the forest environment.

Lastly, it was decided to minimize the factors which might affect the participant during the breaks in-between testing the different conditions. This has been done through a grey environment including a darker grey platform that the participant stands on (see figure 4.3).

Figure 4.3: This figure shows the waiting area of the second experiment.

4.3 Evaluation for Experiment 2

4.3.1 Participants

For this test, 10 participants were recruited. Out of these, 3 participants were female while the remaining 7 were male. All participants were recruited at the CREATE campus of Aalborg University.

4.3.2 Setup & Materials

The same equipment, which was utilized again for Experiment 2 (see *chapter 3, section 3.2.3, Setup & Materials*).

4.3.3 Procedure

As with *Setup & Materials*, the procedure for the second experiment was the same as for the Experiment 1 (see *chapter 3, section 3.2.4, Procedure*). However, the relaxation time in-between the conditions in the experiment has been changed from 5 minutes to 10 minutes. This means that in-between the training session and the first environmental condition, there was a five minute break. The break between the two environmental conditions was increased to being 10 minutes long.

This has been done in order to make sure that there is no carry-over between results. As mentioned in the partial discussion for the first experiment (see *chapter 3, section 3.4, Partial Discussion*), it might be possible that there could be some carry-over between the different results if there has not been enough time for the participants to relax. In the background research (see *chapter 2, Background Research*), it was stated that it was sufficient for the relaxation time in-between conditions to be between 5 and 10 minutes.

Additionally, the participants were also in a VR environment during the breaks.

4.3.4 Problems

Under the experiment, different issues occurred in relation to recording the heart rate data. This resulted in having to redo some of the tests for the experiment, but having to disregard some of the participants, due to time-related issues. Therefore, there were only 8 participants taking part in this experiment.

This was due to the heart rate monitor not recording the measurements correctly. There can be various reasons as to why this happened, including incorrect placement of the heart rate monitor, clothing obstructing the connection between the skin and the heart rate monitor, tightness of the strap securing the heart rate monitor, the amount of charge on the heart rate monitor, as well as faulty equipment. This meant that the heart rate data for 2 of the participants were not complete

and therefore the data of these participants had to be disregarded during the analysis of the results. The heart rate data for each participant was examined between each test.

Because of the problems encountered with the heart rate monitor, the participants generally lacked data under their training session. This data has therefore been excluded from Experiment 2. This should not have any effect on the results, since the stress levels during the training session is not part of the null-hypothesis. For two participants, there was also data missing from the baseline, so those two participants had to be removed from the test. This means that the sample size of this test has been reduced to 8 participants.

4.4 Results

In this section, the results from the experiment will be presented. As mentioned in *Section 2.1*, different variables were looked into, in relation to measuring stress through HRV; LF/HF, HF, pNN50 and r-MSSD. An one-way repeated measures ANOVA was used to look at the data for each variable. Baseline was measured between the training session and the first environment, while Baseline2 was measured between the two environments.

According to t-tests, comparing the order in which participants experienced each condition, did not have any significant effect on the HRV measures of pNN50, r-MSSD, HF and LF/HF.

Figure 4.4 shows a box plot of the LF/HF data. It includes the different conditions made for the one-way repeated measures ANOVA which was used; AvgBaseline, Baseline, Baseline2, Forest, Room. AvgBaseline is the average of the two baselines. Looking at the medians in the boxplot, the highest is Forest. AvgBaseline and Baseline2 is both in-between 2.5 and 5.0 while Baseline is a bit higher, and Room is a bit lower. The ranges vary in size for every condition, Room being the biggest while AvgBaseline and Forest are being the smallest. Out of all five conditions, Forest and Room are the only ones with outliers. Lastly, looking at the lower and upper quartiles for the box plot, the data varies the most for Baseline and Room.

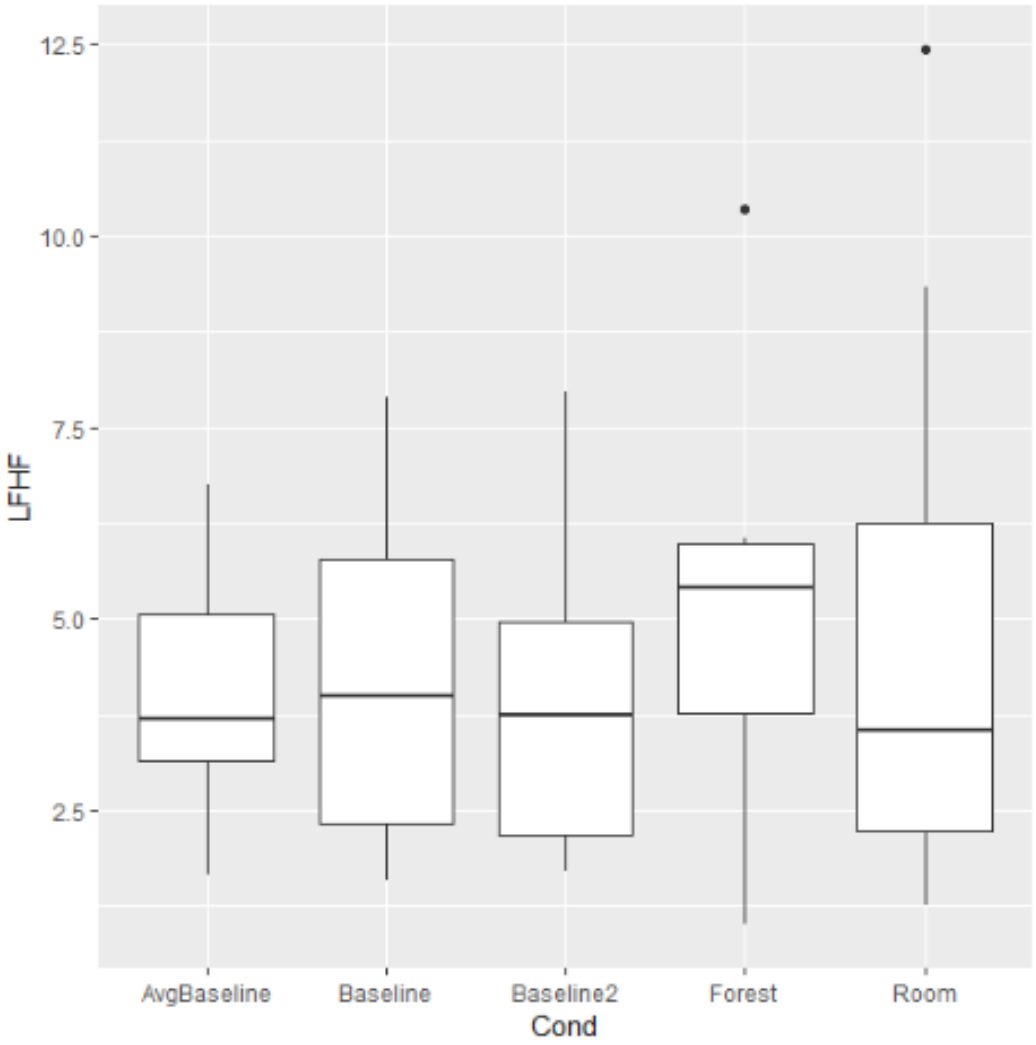

Figure 4.4: In this figure, the boxplot for the LF/HF variable can be seen. The thick horizontal line represents the median of the data, the white boxes represents the upper and lower quartile of the data. The whiskers represents the the biggest and smallest values of the data. The small dots represent the outliers in the data.

Figure 4.5 shows a box plot of the HF data. It includes the different conditions made for the one-way repeated measures ANOVA which was used; AvgBaseline, Baseline, Baseline2, Forest, Room. AvgBaseline is the average of the two baselines. Looking at the medians in the boxplot, the medians for the conditions AvgBaseline, Baseline and Baseline2 are overall higher than the medians for the conditions for the Forest and Room. The ranges vary in size for every condition, Baseline and AvgBaseline being the biggest and Forest being the smallest (however not far from Room). Out of all conditions, AvgBaseline is the only condition without any

outliers. Lastly, looking at the lower and upper quartiles for the box plot, the data overall vary a lot for AvgBaseline and Baseline. Baseline2, Forest and Room have smaller variation in their data.

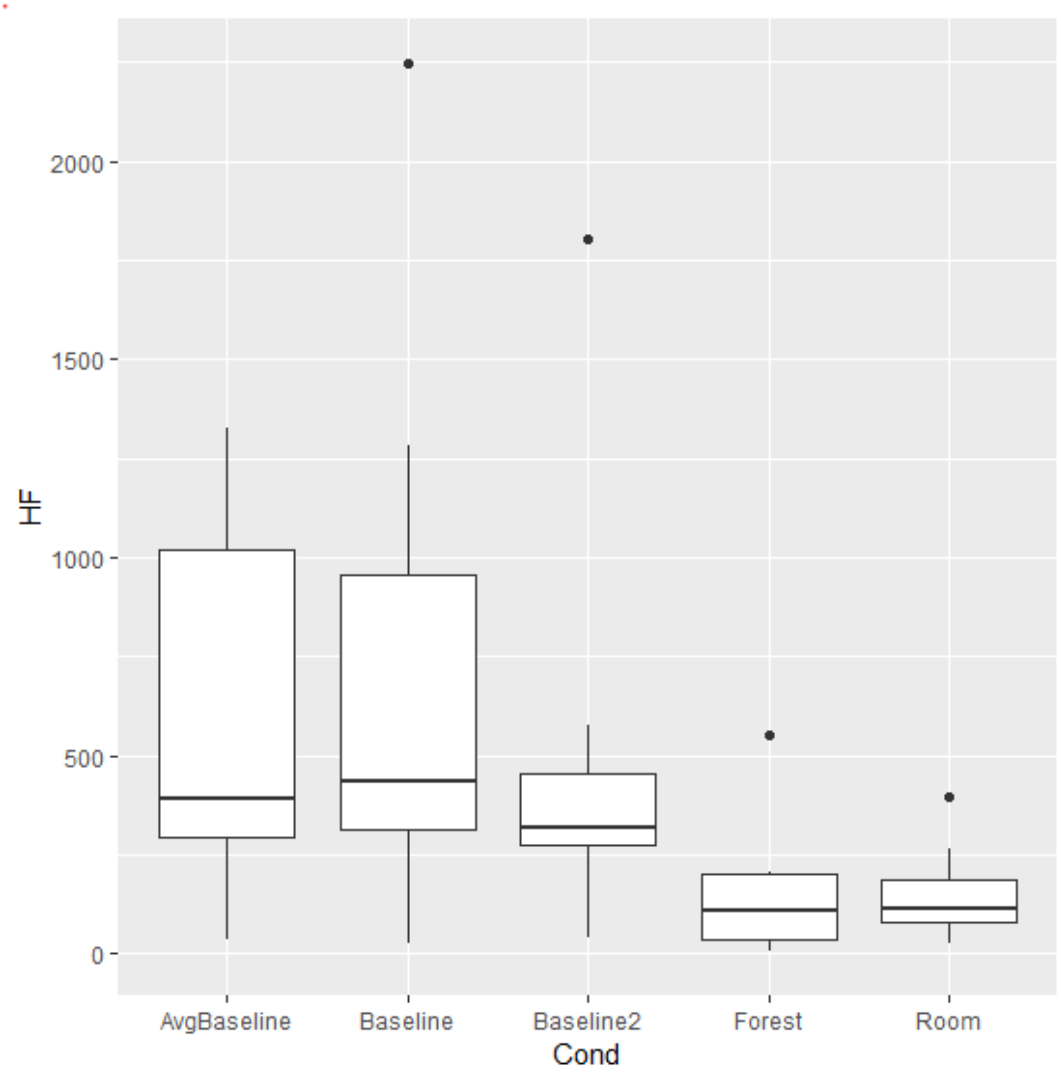

Figure 4.5: In this figure, the boxplot for the HF variable can be seen. The thick horizontal line represents the median of the data, the white boxes represents the upper and lower quartiles of the data. The whiskers represents the the biggest and smallest values of the data. The small dots represent the outliers in the data.

Figure 4.6 shows a box plot of the pNN50 data. It includes the different conditions made for the one-way repeated measures ANOVA which was used; AvgBaseline, Baseline, Baseline2, Forest, Room. AvgBaseline is the average of the two baselines. Looking at the medians in the boxplot, the medians for the conditions AvgBaseline, Baseline and Baseline2 are overall higher than the medians for the

conditions for the Forest and Room. Out of these, Room has a lower median, although not with much difference from Forest. The ranges vary in size for every condition, Baseline2 and AvgBaseline being the biggest and Forest being the smallest. Out of all conditions, Baseline and Forest are the only conditions with outliers. Additionally, Baseline has two outliers, one in the high end and one in the low end. Lastly, looking at the lower and upper quartiles for the box plot, the data overall vary a lot for AvgBaseline, Baseline2 and Room. In regards of this, the data varies slightly less for Baseline and Forest.

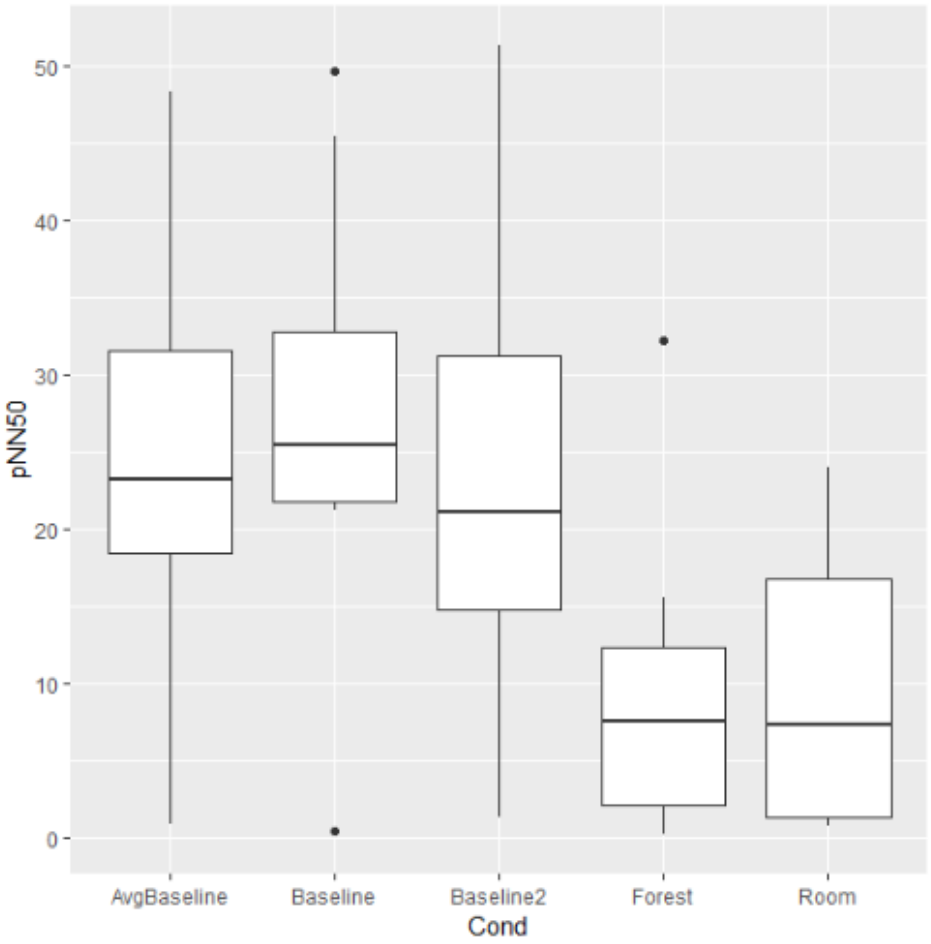

Figure 4.6: In this figure, the boxplot for the pNN50 variable can be seen. The thick horizontal line represents the median of the data, the white boxes represents the upper and lower quartile of the data. The whiskers represents the the biggest and smallest values of the data. The small dots represent the outliers in the data.

Figure 4.7 shows a box plot of the r-MSSD data. It includes the different conditions made for the one-way repeated measures ANOVA which was used; Avg-

Baseline, Baseline, Baseline2, Forest, Room. AvgBaseline is the average of the two baselines. Looking at the medians in the boxplot, the medians for the conditions AvgBaseline, Baseline and Baseline2 are overall higher than the medians for the conditions for the Forest and Room. Out of these, Room has a slightly lower median than Forest. The ranges vary in size for every condition, Baseline being the biggest and Forest and Room being the smallest. Forest and Baseline2 are the only conditions with outliers (one for each). Lastly, looking at the lower and upper quartiles for the box plot, the data overall vary a lot for AvgBaseline and Baseline. The other conditions have smaller variations in their data.

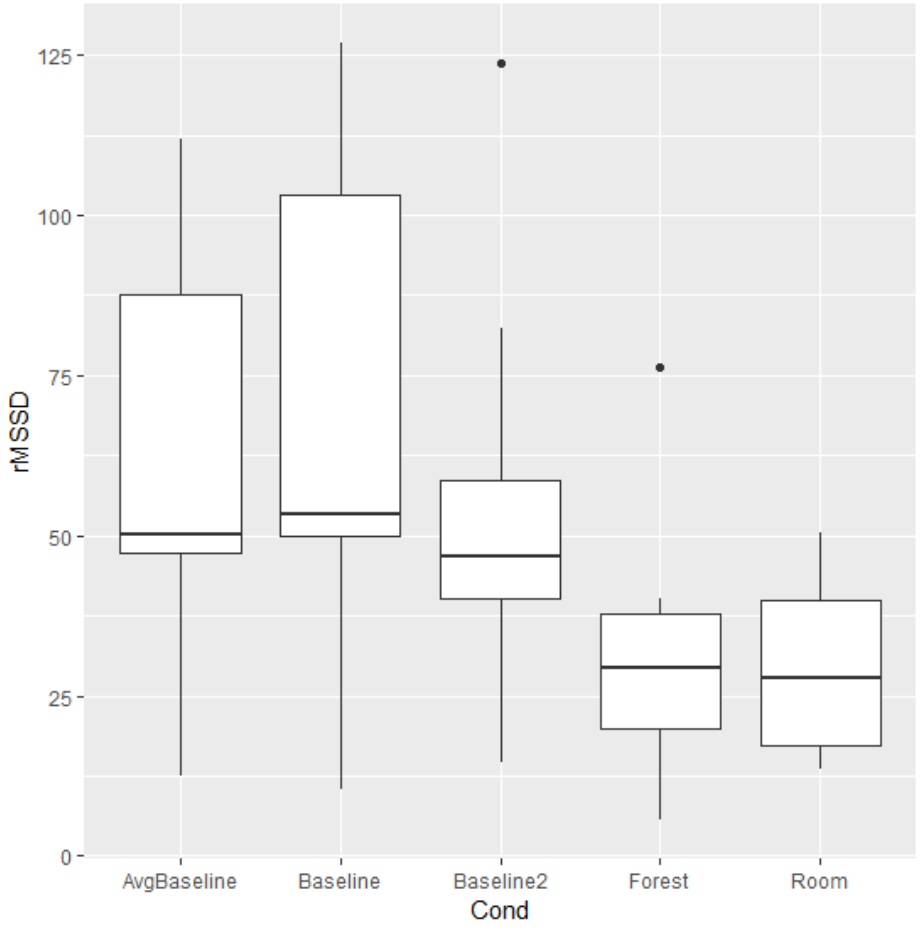

Figure 4.7: In this figure, the boxplot for the r-MSSD variable can be seen. The thick horizontal line represents the median of the data, the white boxes represents the upper and lower quartile of the data. The whiskers represents the the biggest and smallest values of the data. The small dots represent the outliers in the data.

4.4.1 Analysis of Results

In this subsection the presented results will be analysed.

For the LF/HF variable an one-way repeated measures ANOVA showed that the effect of the environment was not significant on LFHF values, $F(4, 35) = 0.301$, $p = 0.87$.

An one-way repeated measures ANOVA also showed that the effect of the environment was not significant on HF values, $f(4, 35) = 2.56$, $p = 0.06$.

In regards to the r-MSSD values, an one-way repeated measures ANOVA showed that the effect of the environment was significant, $F(4, 35) = 2.98$, $p = 0.003$. The post hoc analysis showed that the forest environment ($M = 31.83$, $SD = 21.26$) was significantly different from the baseline ($M = 62.35$, $SD = 32.8$), $z(35) = -3.23$, $p = 0.01$. The simple room environment ($M = 29.25$, $SD = 14.3$) was likewise found to be significantly different from the baseline as well ($M = 62.35$, $SD = 32.8$), $z(35) = -3.5$, $p = 0.004$. However, no significant difference were found between the simple room environment ($M = 29.25$, $SD = 14.3$) and the forest environment ($M = 31.83$, $SD = 21.26$), $z(35) = -0.27$, $p = 0.99$.

Considering the pNN50 values, an one-way repeated measures ANOVA showed that the effect of the environment was significant, $F(4, 35) = 3.3$, $p = 0.021$. This was followed up by a post hoc analysis which showed that the forest environment ($M = 9.77$, $SD = 10.44$) was significantly different from the baseline ($M = 25.76$, $SD = 15.75$), $z(35) = -4.54$, $p < 0.001$. The simple room environment ($M = 9.52$, $SD = 9.13$) was also found to be significantly different from the baseline ($M = 25.76$, $SD = 15.75$) , $z(35) = -4.609$, $p < 0.001$. However, no significant difference were found between the simple room environment ($M = 9.52$, $SD = 9.13$) and the forest environment ($M = 9.77$, $SD = 10.44$), $z(35) = -0.07$, $p = 0.99$, in relation to the pNN50 values.

4.5 Partial Discussion

For this experiment, there are multiple aspects to consider, in terms of reliability and validity, but also the changes which has been made since Experiment 1.

First of all, the procedure of the tests should be considered. The participant needed to be equipped with the heart rate monitor in order for the facilitators to record the data. In order to be able to record the heart rate data, its placement needed to be under their clothing. In relation to this, some people might have been uncomfortable with exposing parts of themselves and made them nervous which might affect the results negatively. Additionally, it might not have been secured correctly for each session which might have resulted in faulty measurements from the monitor.

In terms of the overall procedure, there were breaks in-between each of the

different conditions included in the experiment (training session, forest scene and the simple room scene), in order for them to relax properly. However, it is not certain whether or not the participants have actually relaxed, given that the silence or waiting might have made them uncomfortable. In order to avoid factors outside the VR environments to have an effect, the participants were put in a grey environment during the break.

In terms of the tests, the results might also have been affected by the participants which were recruited. Some of the participants were known to the facilitators which might have produced a form of bias. Secondly, their knowledge and experience with VR might differ. There is a chance that a lack of experience might affect their performance while doing the MIST. Additionally, some participants mentioned that they felt like they were being stressed on purpose. This was not validated by the facilitators, in order to not affect the test results. This might still have affected their performance and reactions during the test, however.

In terms of the equipment used for the test, there might also be different sources of error. During some of the sessions, the controllers stopped reacting momentarily. This was fixed by reconnecting the controllers, but it might have affected the performance during the MIST in one of the VR environments.

Lastly, it should be mentioned that there might not have been a sufficient amount of participants for the experiment. This was a total of 10 participants, which was later reduced to 8, due to faulty measurements. This might therefore affect the reliability of the results.

4.6 Partial Conclusion

For this experiment, it was chosen to improve upon the existing test environments; a simple room and a forest scene. These changes included a better incorporation of the MIST implementation in the different environments. Additionally, for the forest scene, there were some changes in order to make it seem more realistic and natural. Lastly, it was decided to put the participants into VR during the breaks in-between the different conditions (including Training). This was done in order to minimize the factors that might affect the participant while outside of VR.

Looking at the analysis of results in *section 4.4.1*, the null-hypothesis could not be refuted, as there were no significant difference in the stress levels between the two environments.

Chapter 5

Experiment 3

In this chapter, the third experiment will be explained in further detail.

5.1 Purpose of the Experiment

When looking at the findings of Experiment 2, the null-hypothesis which had been set up could not be refuted. However, as in Experiment 1 (see *chapter 3, Experiment 1*), it could be confirmed that the implemented MIST did stress the participants.

As mentioned in *chapter 2, Background Research* a person should be more relaxed when being in a natural setting compared to an urban one. It was also stated that a higher amount of nature elements could decrease stress even further. Therefore it could be hypothesised that the nature scene in Experiment 2 was not immersive enough to make the participants feel like they were in nature. Therefore it has been decided to look into how to improve the immersion of a virtual environment, in order to help the participants feel like they are in the forest or simple room.

According to S. Huiberts[5] adding audio compatible with the scenery of a game or any other digital environment can be a great help to build a users immersion in a scene. This audio could e.g. be a non-diegetic sound to help the user get a feeling of the environment, using sound to alert a player to other creatures in a scene not yet in view, etc.

Because of this it has been decided to implement non-diegetic sound into each of the environments to help the participants to feel immersed in the environments

As this experiment is based on the same assumptions as Experiment 1 and Experiment 2, it was decided to use the same null-hypothesis as those experiments:

"A forest scene is not better at decreasing stress in participants when performing a mental stress task."

5.2 Design & Implementation

For the third experiment, it was decided to implement sound for a better experience and immersion.

The soundtrack "Evening in the Forest" made by *reinsamba* from Freesound[12] was used in order to create an ambient background sound for the forest scene. This soundtrack mainly consists of birds singing, as well as cicadas. Changes included prolonging the soundtrack to 5.09 minutes instead of 2.36 minutes. This was done using the audio editing software Audacity.

In regards to changes done to the forest scene, it was also decided to add a skybox to the scene in order to improve the aesthetics of the scene (see figure 5.1).

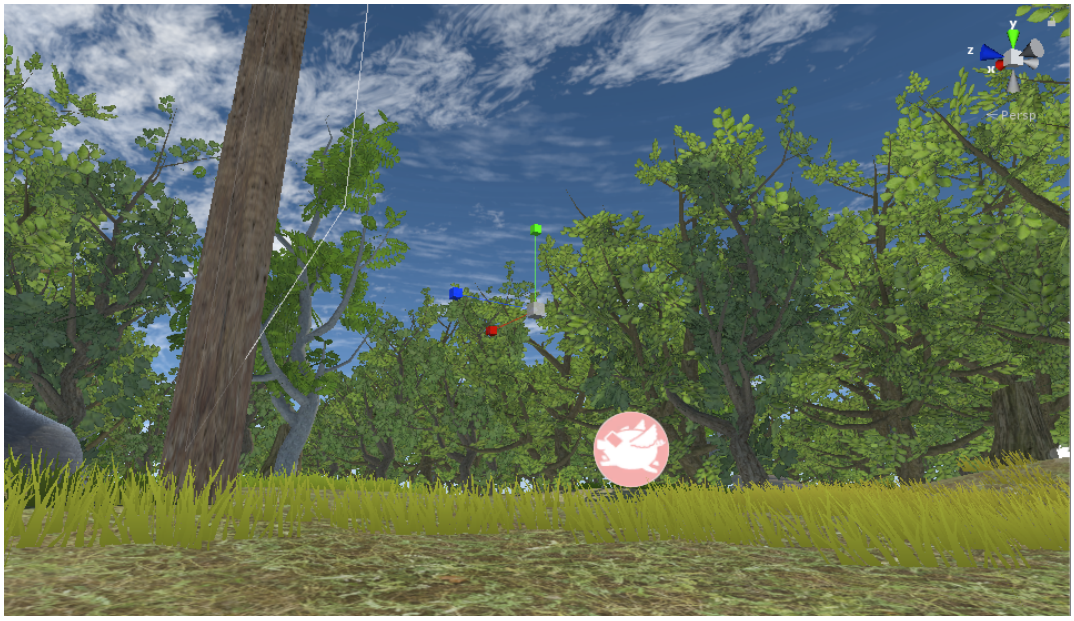

Figure 5.1: This figure shows the skybox which was added to the third version of the forest environment.

For the simple room scene, "11 minutes of city sounds" was used. This soundtrack is from Freesound and made by *Niedec*[10]. It consists of traffic noise, birds singing, people talking and dogs barking. Changes included editing the bass and tremble, as well as the volume on the track, in order to make it sound more distant and distorted, as the sound is supposed to come from behind a wall. This was done using the audio editing software Audacity.

5.3 Evaluation for Experiment 3

5.3.1 Participants

For this experiment, 10 participants were recruited. Of these participants, 6 were male and 4 were female. All participants were recruited at the CREATE campus of Aalborg University.

5.3.2 Setup & Materials

For the third experiment, the same equipment as was used for the first and second experiment, was utilized again (see *chapter 3, section 3.2.3, Setup & Materials*). Additionally, a speaker was also used to present the newly added audio to the participants.

5.3.3 Procedure

The procedure for the third experiment was the same as for the first and second experiment (see *chapter 3, section 3.2.4, Procedure*).

5.3.4 Problems

In regards to problems occurring during the experiment, these concerned the heart rate monitor and recording the heart rate data. This not only meant that some tests had to be redone, but also exclusion of two participants during the analysis of the results. Therefore there were only 8 valid participants for this experiment. This was due to the heart rate monitor not recording the heart rate measurements correctly. There can be various reasons as to why this occurred, including incorrect placement of the heart rate monitor, clothing which can have caused an obstruction of the connection between the skin and the heart rate monitor, tightness of the strap securing the heart rate monitor, the amount of charge on the heart rate monitor as well as faulty equipment. There were one participants which was deemed to have faulty measurements for this experiment, and therefore excluded. In order to make sure that there is an equal amount of participants on starting with each condition, a random participant was excluded from the experiment.

Another element which could be a problem for the validity of the results, is that during testing audio from an adjacent room could sometimes be heard. This could mean that the participant lost some of the immersion the newly added sound was trying to create.

5.4 Results

In this section, the results from the test will be presented. As with the two other experiments done for this project, the following variables were looked into, in relation measuring stress through HRV; LF/HF, HF, pNN50 and r-MSSD. An one-way repeated measures ANOVA was used to look at the data for each variable. Baseline was measured between the training session and the first environment, while Baseline2 was measured between the two environments.

According to t-tests, comparing the order in which participants experienced each condition, order did not have any significant effect on the HRV measures of LF/HF, HF and pNN50. However, for the r-MSSD measurements, order showed a significant difference for the simple room environment (Forest first mean = 26.14, Room first mean = 18.4, p = 0.036). Meaning that participants that experienced the Room scenario first had a significantly lower r-MSSD value, while in the Room environment, than those experiencing the Forest first. This suggests that there might have been some carry over from the forest environment.

Figure 5.2 shows a box plot of the LF/HF data. It includes the different conditions made for the one-way repeated measures ANOVA which was used; AvgBaseline, Baseline, Baseline2, Forest, Room and Training. AvgBaseline is the average of the two baselines.

Looking at the medians in the boxplot, the highest is Training (the only median with a value above 5.0). The lowest is Baseline2. The ranges vary in size for every condition, Forest being the biggest. The range of AvgBaseline varies the least. Apart from Training, all conditions include one outlier. Additionally, AvgBaseline includes two outliers, one in the high end and one in the low end. Lastly, looking at the lower and upper quartiles for the box plot, the data varies the most for Forest and Training, while varying the least for AvgBaseline and Room.

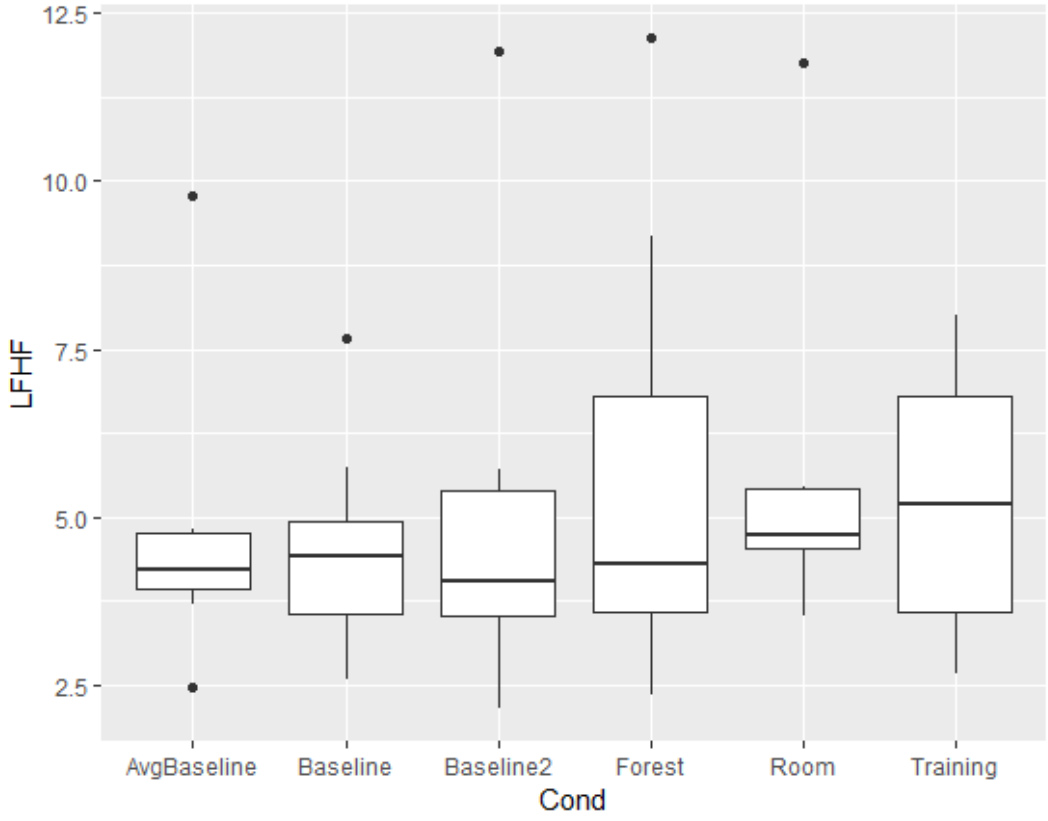

Figure 5.2: In this figure, the boxplot for the LF/HF variable can be seen. The thick horizontal line represents the median of the data, the white boxes represents the upper and lower quartile of the data. The whiskers represents the the biggest and smallest values of the data. The small dots represent the outliers in the data.

Figure 5.3 shows a box plot of the HF data. It includes the different conditions made for the one-way repeated measures ANOVA; AvgBaseline, Baseline, Baseline2, Forest, Room and Training. AvgBaseline is the average of the two baselines.

Looking at the medians in the boxplot, AvgBaseline, Baseline and Baseline2 are overall the highest. The range vary the most for Baseline2 and the least for Forest. Apart from Forest and Room, all conditions include one outlier. Lastly, looking at the lower and upper quartiles for the box plot, the data varies the most for AvgBaseline and Baseline2, while varying the least for Forest (although there is not a lot of difference from Room and Training).

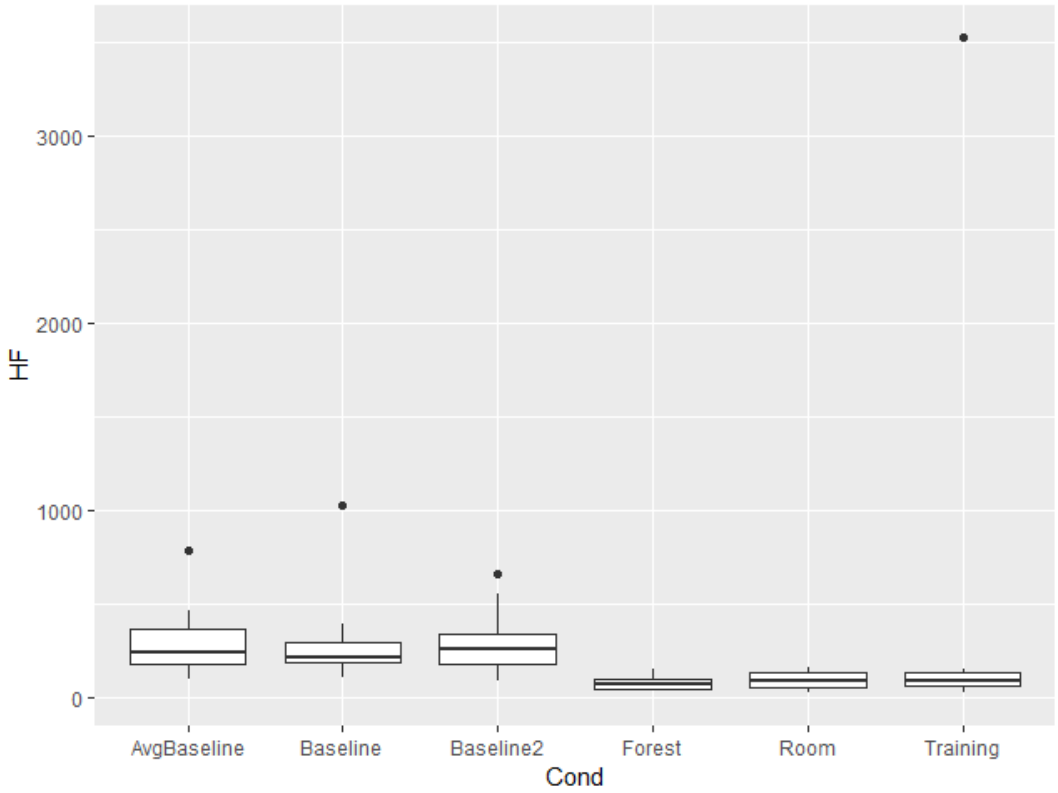

Figure 5.3: In this figure, the boxplot for the HF variable can be seen. The thick horizontal line represents the median of the data, the white boxes represents the upper and lower quartile of the data. The whiskers represents the the biggest and smallest values of the data. The small dots represent the outliers in the data.

Figure 5.4 shows a box plot of the pNN50 data. It includes the different conditions made for the one-way repeated measures ANOVA; AvgBaseline, Baseline, Baseline2, Forest, Room and Training. AvgBaseline is the average of the two baselines.

Looking at the medians in the boxplot, Baseline2 is the highest while Forest has the lowest median. The range vary the most for the three baselines and the least for Forest. Forest and Room are the only conditions which each include an outlier. Lastly, looking at the lower and upper quartiles for the box plot, the data varies the most for Baseline2 and while varying the least for Room.

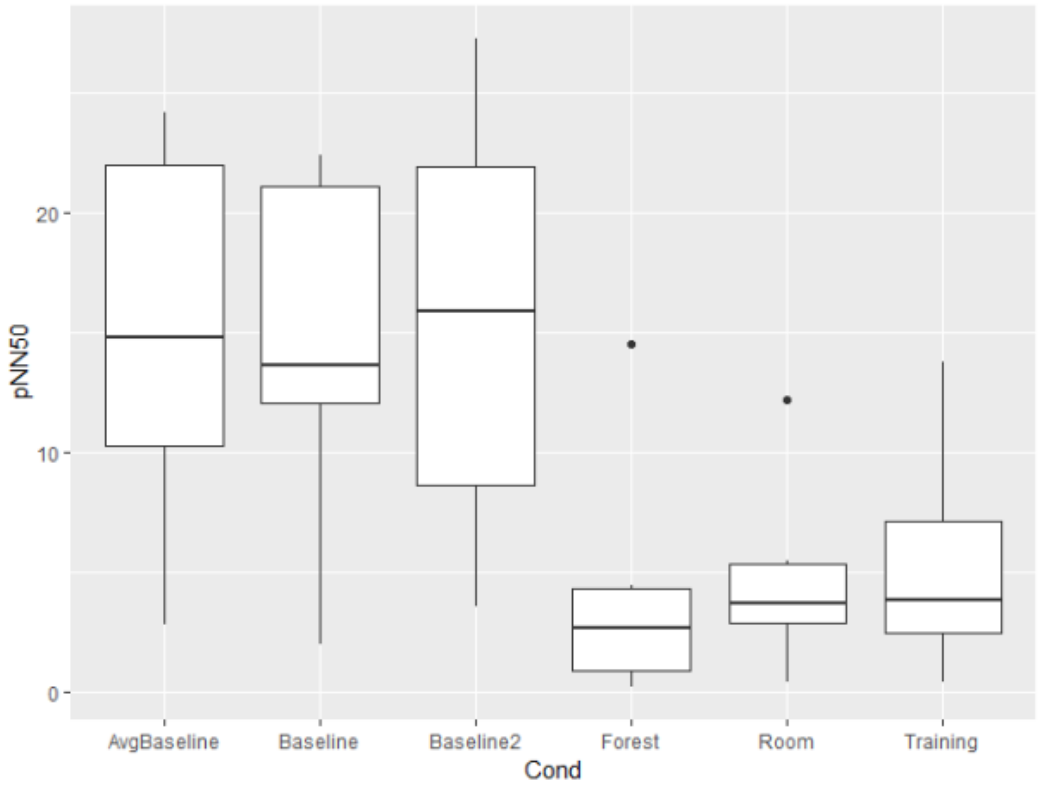

Figure 5.4: In this figure, the boxplot for the pNN50 variable can be seen. The thick horizontal line represents the median of the data, the white boxes represents the upper and lower quartile of the data. The whiskers represents the the biggest and smallest values of the data. The small dots represent the outliers in the data.

Figure 5.5 shows a box plot of the rMSSD data. It includes the different conditions made for the one-way repeated measures ANOVA; AvgBaseline, Baseline, Baseline2, Forest, Room and Training. AvgBaseline is the average of the two baselines.

Looking at the medians in the boxplot, the baseline conditions have the highest medians, the others being generally lower and similar to each other. The range vary the most for AvgBaseline and the least for Room (although there is not a big difference from this condition and Forest and Training). Apart from AvgBaseline and Room, all conditions each include an outlier. Lastly, looking at the lower and upper quartiles for the box plot, the data varies the most for AvgBaseline and while varying the least for Room.

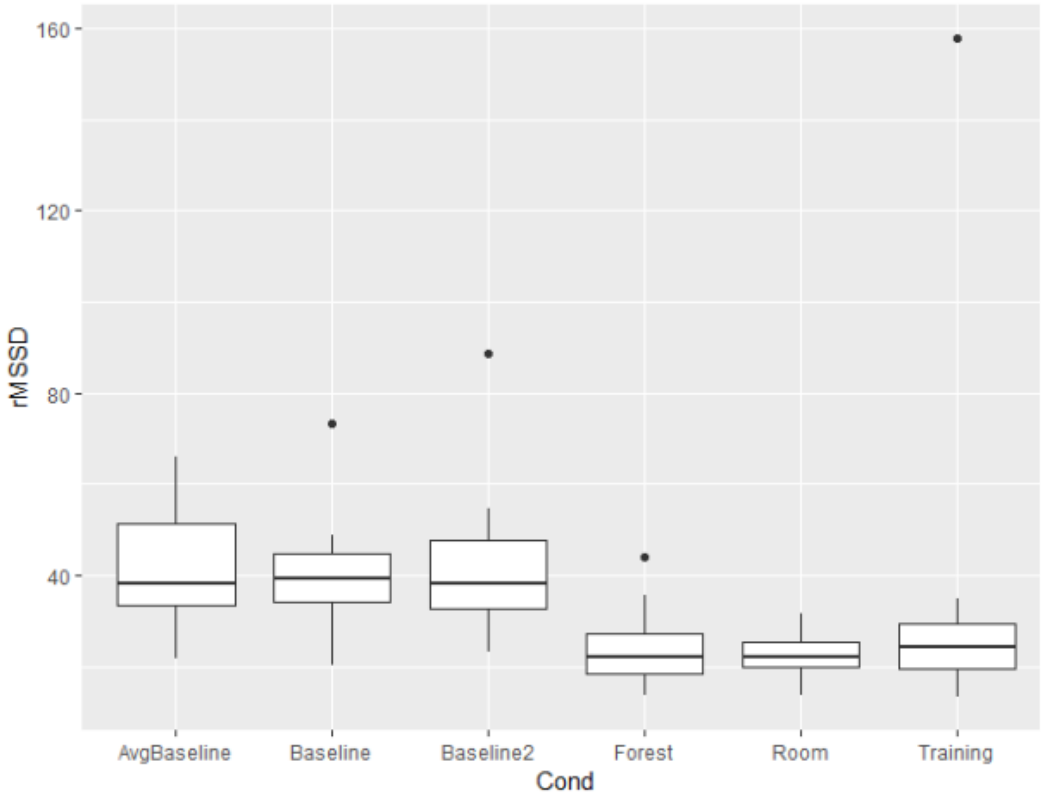

Figure 5.5: In this figure, the boxplot for the rMSSD variable can be seen. The thick horizontal line represents the median of the data, the white boxes represents the upper and lower quartile of the data. The whiskers represents the the biggest and smallest values of the data. The small dots represent the outliers in the data.

5.4.1 Analysis of Results

An one-way repeated measures ANOVA showed that the effect of the environment was not significant on LFHF values, $F(5, 42) = 0.26$, $p = 0.93$.

An one-way repeated measures ANOVA showed that the effect of the environment was not significant on HF values, $F(5, 42) = 0.77$, $p = 0.58$.

An one-way repeated measures ANOVA showed that the effect of the environment was significant on r-MSSD values, $F(5, 42) = 7.1$, $p > 0.0001$. Post hoc analysis using the Tukey HSD method showed that the Forest environment ($M = 22.22$, $SD = 10.17$) was significantly different from the baseline ($M = 38.35$, $SD = 15.54$), $z(42) = -6.31$, $p < 0.001$. The Room environment ($M = 22.17$, $SD = 5.59$) was likewise found to be significantly different from the baseline ($M = 38.35$, $SD = 15.54$), $z(42) = -5.94$, $p < 0.001$. No significant difference were found between the Room environment ($M = 22.17$, $SD = 5.59$) or the Forest environment ($M = 22.22$, $SD = 10.17$), $z(42) = 0.37$, $p = 0.99$.

An one-way repeated measures ANOVA showed that the effect of the environment was not significant on pNN50 values, $F(5, 42) = 1.36$, $p = 0.259$.

5.5 Partial Discussion

For this experiment, there are multiple aspects to consider, in terms of reliability and validity, but also the changes which has been made since Experiment 2.

One thing to be considered is the procedure of the experiment. Because of how the heart rate monitor records data, it is important for the heart rate monitor to have unrestricted access to the participants' skin. Because of this, the heart rate monitor had to be placed against skin under a participants' clothes. Furthermore, in order to attempt to make sure that the heart rate monitor was recording data, the facilitator would have to look at the heart rate monitor on the participant until a red light would blink. This could have made the participants uncomfortable, which could have affected the results negatively. Furthermore, the heart rate monitor might not have been secured correctly for each session which might have resulted in faulty measurements from the monitor.

Another aspect to consider is the recruited participants. Some of the participants which were recruited knew the facilitators. This could mean that these participants were less nervous during the tasks than the rest of the participants. Furthermore, another participant bias could be the difference in experience with a VR environment which could mean that some participants were more excited or nervous than other because of the technology.

In terms of the overall procedure, there were breaks in-between each of the different conditions included in the experiment (training session, forest scene and the simple room scene), in order for them to relax properly. However, it is not certain whether or not the participants have actually relaxed, given that the silence or waiting might have made them uncomfortable. In order to avoid factors outside the VR environments to have an effect, the participants were put in a grey environment during the break.

In terms of the equipment used for the test, there might also be different sources of error. During some of the sessions, the controllers stopped reacting momentarily. This was fixed by reconnecting the controllers, but it might have affected the performance during the MIST in one of the VR environments.

It should furthermore be mentioned that there might not have been a sufficient amount of participants for the experiment. This was a total of 10 participants, which was later reduced to 8, due to faulty measurements. This might therefore affect the reliability of the results.

Lastly, it should be mentioned that during some test the room adjacent to the testing area sometimes had faint music playing. This could have affected how the participants perceived the audio in the two environments, which could have effect

the validity of the results.

5.6 Partial Conclusion

For this experiment, it was chosen to improve upon the existing test environments by adding diegetic audio to each scene, as well as a skybox to the forest scene. This was done in order to create a better immersion.

Looking at the analysis of results in *section 5.4.1, Analysis of Results*, the null-hypothesis could not be refuted, as there were no significant difference in the stress levels between the two environments. However, the analysis of results did show that there were less variables showing a significant difference between the base-lines and the different conditions, compared to the other experiments. This could indicate that the participants were less stressed while having diegetic audio in the environments as well.

Chapter 6

Discussion

In this chapter, the interpreted results of the experiments will be discussed, along with reflections on the procedure of the experiments.

6.1 Interpretation of Results

Looking at *section 3.3.1, Analysis of Results* for the first experiment, it was not possible to refute the null-hypothesis with the calculated p-values. There were no significant difference between the two VR environments. However, it was found that the MIST did in fact stress the participants.

For the second experiment, it was not possible to refute the null-hypothesis either (see *section 4.4.1, Analysis of Results*). There was no significant difference between the VR environments, despite of the changes which had been made to improve them. It was however still possible to see a difference between the baselines and the different environments which indicates that the participants were in fact stressed by the implemented MIST.

For the third and last experiment, it was not possible to refute the null-hypothesis (see *section 5.4.1, Analysis of Results*). There was no significant difference between the different VR environments, despite of the incorporated diegetic audio. However, less variables showed a significant difference between the baselines and the conditions, compared to the other experiments. This indicates that the participants might have been less stressed which points towards audio being effective in relation to deducing stress.

6.2 Relation to Other Works

In the state of art in relation to deducing stress (see *section 2.3 What is the state of the art to induce relaxation?*), different projects using environmental factors to deduce stress were looked into. These projects generally looked into how nature elements

could affect relaxation. Jiang et al.[6] found that having more nature elements in a scene would help to reduce stress. Additionally, these projects did their experiments in real-world surroundings. The aim for this project was therefore to test these environments in a virtual setting instead in order to eliminate the amount of factors affecting the results. Furthermore, the aim was to find out what specific elements had an effect in relation to deducing stress.

However, with this project, it was not possible to back these findings. This indicates that while environments might have an effect on relaxation when doing nothing, or perhaps decreasing the time needed for relaxation. It appears that the environment does not affect stress levels when performing cognitive tasks.

6.3 Impact for practitioners

As mentioned in *section 6.1, Interpretations of Results*, it was not possible refute either of the null-hypotheses in the different experiments. This indicates that environments should not be able to affect stress either positively or negatively. However, during these experiments, it was found that the implemented MIST did stress the participants and that audio might have an effect in relation to reduce stress overall, during a mental stress task.

6.4 Limitations

Looking at the limitations for this project, there were some issues which are worth mentioning.

The problems which have occurred during the experiment should be mentioned. These concerned the heart rate monitor and its recording of heart rate data. At some points, the data was not recorded properly. There can be various reasons as to why this happened, including incorrect placement of the heart rate monitor, clothing which can have caused an obstruction of the connection between the skin and the heart rate monitor, tightness of the strap securing the heart rate monitor, the amount of charge on the heart rate monitor as well as faulty equipment. This resulted in some tests having to be redone, as well as participants which had to be excluded during the process of analysing the data. This meant less data which affects the reliability of the results.

Another limitation to consider is the number of participants in each experiment. One experiment had 10 while the other two only had 8. This could have affected the reliability and validity of the results since it would have been preferable to have more participants. Another aspect in regards to the participants is that some of the participants in all three experiments were known to the facilitators which might have affected the results.

Additionally, there might have been a difference in the experience with virtual reality among the participants. Some had worked with it previously, others had not. This might affect the validity of the results since it might have affected their experience during the tests, in terms of distracting them the task at hand or stressing them even more.

During the third and last experiment, diegetic audio was implemented for a better immersion. However, during the time of the experiment, there were other sources of audio which could be heard through the walls of the room in which the experiment took place. This might have affected the experience and immersion in the VR environments and therefore also the validity and reliability of the results.

Additionally, a last limitation to consider is the controller used for the VR experience. When the participant had to do the MIST the controller would sometimes get stuck between two numbers for a short time. This issue could be fixed by reconnecting the controller. This can however have affected the participants' ability to do the MIST and have stressed them more than the participants who did not encounter this problem. This can have affected the reliability and validity of the results.

6.5 Future Works

If future work were to be done on this project, several issues should be considered.

First of all, in terms of the reliability of the results, more participants and data would be ideal.

Apart from this, it might be ideal to work with another heart rate monitor, in order to get better and more precise measurements. Additionally, taking samples of saliva could give additional recordings of stress levels.

Chapter 7

Conclusion

For this project, it was found that several people suffer from or experience symptoms of severe stress. Therefore, the aim was to find a way to reduce stress.

Through the background research, different ways of measuring stress were looked into, as well as what methods are already used in relation to reducing stress. It was found that nature environments are commonly used as a way to induce relaxation, while also reducing stress. From this, the following problem statement was made:

"What elements in a virtual environment are responsible for inducing relaxation?"

Three experiments were conducted, in which two different virtual environments (a simple room and a forest scene) were tested against each other with an implemented MIST. For the first experiment, no significant difference were found between the simple room and the forest scene. The null-hypothesis;

"A forest scene is not better at decreasing stress in participants when performing a mental stress task."

could therefore not be refuted. As there was a difference between the baselines and the environments, it could however be found that the MIST did in fact stress people. As the background research showed that nature should have an effect in relation to reducing stress, it was therefore decided to improve upon the environments and repeat the experiment.

For the second experiment, no significant difference between the virtual environments were found. Therefore, the null-hypothesis (the same as for the first experiment) could not be refuted during this experiment either.

For the third and last experiment, audio was implemented in order to create a better immersion within the scenes. The null-hypothesis could not be refuted with the calculated p-values. However, less variables showed a significant difference between the baselines and the environments, compared to the previous experiments. This might indicate that audio could have an effect in relation to reducing stress.

From these three experiments, in which the null-hypotheses could not be refuted, it is not certain that nature has an effect in relation to reducing stress. However, this might have to do with the immersion of the virtual environments. Additionally, there might be indications towards audio being effective in reducing stress, and therefore further research on this topic are to be encouraged.

Bibliography

[1] Katja Borchert. *MONTREAL IMAGING STRESS TEST (no port info)*. https://www.millisecond.com/download/library/v5/montrealstresstest/montrealstresstest.manual, Accessed 20th of March 2019.

[2] Gregory N Bratman et al. "The benefits of nature experience: Improved affect and cognition". In: *Landscape and Urban Planning* 138 (2015). ,https://innerlijkefocus.nl/wp-content/uploads/Benefits-of-nature-experience-improved-affect-and-cognition.pdf, Accessed 21st of February 2019, pp. 41–50.

[3] Katarina Dedovic et al. "The Montreal Imaging Stress Task: using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain". In: *Journal of Psychiatry and Neuroscience* 30.5 (2005). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1197276/, Accessed 20th of February 2019, p. 319.

[4] HealthLink BC. *Stress Management: Breathing Exercises for Relaxation*. https://www.healthlinkbc.ca/health-topics/uz2255, Accessed 21st of February 2019.

[5] Sander Huiberts. "Captivating sound the role of audio for immersion in computer games". https://download.captivatingsound.com/Sander_Huiberts_CaptivatingSound.pdf, Accessed April 30th 2019. PhD thesis. University of Portsmouth, 2010.

[6] Bin Jiang et al. "A dose-response curve describing the relationship between urban tree cover density and self-reported stress recovery". In: *Environment and behavior* 48.4 (2016). https://journals-sagepub-com.zorac.aub.aau.dk/doi/pdf/10.1177/0013916514552321, Accessed 21st of February 2019, pp. 607–629.

[7] Hye-Geum Kim et al. "Stress and heart rate variability: A meta-analysis and review of the literature". In: *Psychiatry investigation* 15.3 (2018). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5900369/pdf/pi-2017-08-17.pdf, Accessed February 25th 2019, p. 235.

[8] Clemens Kirschbaum, Karl-Martin Pirke, and Dirk H Hellhammer. "The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting". In: *Neuropsychobiology* 28.1-2 (1993). https://www.ncbi.nlm.nih.gov/pubmed/8255414, Accessed 20th of February 2019, pp. 76–81.

[9] I. Mahalil et al. "Virtual reality-based technique for stress therapy". In: *2014 4th International Conference on Engineering Technology and Technopreneuship (ICE2T)*. https://ieeexplore.ieee.org/document/7006265; Accessed 21st of February 2019. 2014, pp. 295–300. DOI: 10.1109/ICE2T.2014.7006265.

[10] Niedec. *11 minutes of city sounds*. https://freesound.org/people/Niedec/sounds/250270/, Accessed April 27th 2019.

[11] Stephen W Porges. "Cardiac vagal tone: a physiological index of stress". In: *Neuroscience & Biobehavioral Reviews* 19.2 (1995). https://www.sciencedirect.com/science/article/pii/014976349400066A, Accessed February 23rd 2019, pp. 225–233.

[12] reinsamba. *evening in the forest.wav*. https://freesound.org/s/18765/I, Accessed April 26th 2019.

[13] Florian Soyka et al. "Enhancing stress management techniques using virtual reality". In: https://www.researchgate.net/publication/304533181_Enhancing_stress_management_techniques_using_virtual_reality, Accessed 21st of February 2019. July 2016, pp. 85–88. DOI: 10.1145/2931002.2931017.

[14] Stress foreningen. *Stress og statistik*. http://www.stressforeningen.dk/stress-og-statistik/ Accessed 19th of February 2019.

[15] Sundheds Styrelsen. *LANGVARIG STRESS Aktuel viden og forslag til stressforebyggelse RÅDGIVNING TIL ALMEN PRAKSIS*. https://www.sst.dk/~/media/45A1292B2CF643BFB2DD80765188AD85.ashx, Accessed 20th of February 2019.

[16] Sangthong Terathongkum and Rita H. Pickler. "Relationships among heart rate variability, hypertension, and relaxation techniques". In: *Journal of Vascular Nursing* 22.3 (2004). https://www.ncbi.nlm.nih.gov/pubmed/15371972, Accessed 19th of February 2019, pp. 78 –82. ISSN: 1062-0303. DOI: https://doi.org/10.1016/j.jvn.2004.06.003. URL: http://www.sciencedirect.com/science/article/pii/S1062030304000986.

[17] The American Institute of Stress. *Self-Assessment*. https://www.stress.org/self-assessment, Accessed April 23rd 2019.

[18] K. D. Thoondee and A. Oikonomou. “Using virtual reality to reduce stress at work”. In: *2017 Computing Conference*. https://ieeexplore.ieee.org/document/8252142, Accessed 21st of February 2019. 2017, pp. 492–499. DOI: 10.1109/SAI.2017.8252142.

[19] Bin Yu et al. “A study on user acceptance of different auditory content for relaxation”. In: *Proceedings of the Audio Mostly 2016*. https://dl.acm.org/citation.cfm?id=2986418, Accessed February 25th 2019. ACM. 2016, pp. 69–76.