REAL-TIME MONITORING OF PHYSIOLOGICAL AROUSAL: A ‘SERIOUS’ OPTION?

A RESEARCH ABOUT ELECTRODERMAL ACTIVITY, HEART RATE, AND HEART RATE VARIABILITY DURING A SERIOUS GAME.

Joyce Leussink

Master Thesis

Faculty of Behavioural Sciences
Department of Cognitive Psychology & Ergonomics (CPE)
University of Twente

Enschede, 19th of May 2015
Subject: Master Thesis HFE

Supervisors:
Dr. Matthijs Noordzij, University of Twente
Suzanne Vosslamber, MSc, University of Twente
Ing. Thomas de Groot, T-Xchange

UNIVERSITY OF TWENTE. T-Xchange

"The mind is not a book, to be opened at will and examined at leisure. Thoughts are not etched on the inside of skulls, to be perused by an invader. The mind is a complex and many-layered thing."

J.K. Rowling, Harry Potter and the Order of the Phoenix

"Doubt is the origin of wisdom."

René Descartes

Abstract

In this thesis we want to investigate if it is useful to control physiological measures during a serious game in real-time. The reason for this is that the external stakeholder, T-Xchange, hopes that they can adapt a serious game based on the arousal of the players. To study this we used one slow and one fast version of a game and looked at EDA, HR, and HRV for an indication about the physiological arousal. A survey was also used to test the subjective stress of the participants. It was expected that the fast version of the game would elicit more physiological arousal than the slow version. 43 participants, randomly divided over the two versions, participated in this experiment. The differences between the two versions were tested. It was also tested if the physiological measures correlated with each other and if the self-reports correlated with these physiological measures. No significant differences in the EDA and HRV were found. However, the HR in the baseline was significantly higher than during the game and the HR in the slow version was higher than in the fast version. Furthermore, no correlations between the physiological measures were found. No relations were found between the self-reports and physiological measures either. It can be concluded that real-time information about physiological arousal has to be interpreted with caution. We recommend T-Xchange to not adapt their game based on physiological measures, but use a serious game that has some short movies in it that will make the player feel like he or she is participating in a realistic experience. This will increase the transfer of learned information and players will use this information better in real situations.

Acknowledgments

First of all I would like to thank Matthijs Noordzij and Suzanne Vosslamber for their inspiration, good support, and feedback. Secondly, I would like to thank Thomas de Groot for his practical advice about the game that I used and his feedback. Furthermore, I would like to thank Johan de Heer for the inspirational talks we had, my colleagues at T-Xchange for their help and laughter when I needed it and of course the participants that volunteered in my experiment. Without them this thesis could not be written.

Special thanks go out to my family, boyfriend, and friends who always believed in me and listened to me when I was talking to them about my thesis.

Table of Contents

1. Introduction

1.1 Serious games

These days, computer games are mostly seen as a source of entertainment. However, the last few years, computer games have been made that serve another purpose: learning. These games are called ‘Serious games’ (Susi, Johannesson, & Backlund, 2007).

According to Zyda (2005), a serious game is “A mental contest, played with a computer in accordance with specific rules, that uses entertainment to further government or corporate training, education, health, public policy, and strategic communication objectives” (p. 26). He wrote that serious games have more than just story, art, and software. They also involve pedagogy: activities that educate or instruct and thereby give knowledge or teach a skill. Although Zyda (2005, p. 26) thought that the most important part of a serious game was the story and thus the entertainment, Miller, Chang, Wang, Beier, and Klisch (2011) defined serious games as “games primarily focused on education rather than entertainment” (Miller et al., 2011, p. 1425). Besides that, they state that serious games are videogames with a useful purpose for training, education, knowledge acquisition, and skill development.

Some people think that the term ‘serious game’ is a contradiction between the terms ‘serious’ and ‘game’, because most known games are only for fun and relaxation. Although the term ‘serious’ describes what the purpose of the game is and why it was created, it does not say anything about the content of the game (Susi et al., 2007). Serious games are used in a wide variety of fields, like business, industry, marketing, healthcare, government, and education (Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012).

1.2 Cognitive workload

When people are learning, thinking, or reasoning they will have to put some cognitive effort in their tasks. This perceived effort is also called *cognitive load* or *cognitive workload* (Shi, Ruiz, Taib, Choi, & Chen, 2007), and is defined as the amount of mental energy that is required to process a given amount of information (Ang, Zaphiris, & Mahmood, 2007). Cognitive load can be influenced by increasing the level of difficulty of a task or by adding a secondary task. Especially information that is presented simultaneously creates heavy cognitive loads upon the participant because our working memory has problems with understanding and processing this (Ang et al., 2007). The less cognitive load a person has to handle, the easier and faster a task can be completed.

Gaming also requires cognitive load. Many games require players to think, reason, and to handle different tasks at the same time. To analyse this, different studies about cognitive load in the field of human computer interaction have been performed. For example, Ang et al. (2007) studied what kinds of cognitive load exists when a massively multiplayer online role playing game (MMORPG) is played. They also investigated what the effect of cognitive load on the players' performance and engagement in the game was. Their results revealed that cognitive overload exists during game playing. This cognitive overload caused serious problems, even for expert players. However, players seemed to develop strategies to overcome these problems, such as using keyboard shortcuts to save time. Besides this, they found that some forms of cognitive load are desirable for making the game challenging.

Cognitive load can be measured in three different ways: by using performance measures, by using rating scales, or by using physiological measures (Ikehara & Crosby, 2005; Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Performance measures rely on the time taken for a task or the percentage of the task that is performed correctly. However, tasks are often different from each other, so the performance measure should always be adapted to a new task and that takes time. Using rating scales is a subjective method for measuring the cognitive load. It is often used after a task has been completed. It gives insight in the user's perception of the task but does not give real-time information about the cognitive load during this task. The last method is using physiological measures. This can be used during a task without disturbing the user and gives real-time information about physiological changes that can be related to *physiological arousal* (Ikehara & Crosby, 2005).

Various studies examined physiological arousal in combination with high or low cognitive load. Most of these studies were in the field of driving or aviation. For example, Brookhuis and De Waard (2001), who studied heart rate while driving and discovered that heart rate increases with heightened task demand. In their research, this heightened task demand was created by entering a traffic circle. In another study it was found that cardiac measures are sensitive to attentional costs and can be used in research about performance measures (Lenneman & Backs, 2009).

Although there are studies about cognitive workload and gaming, and cognitive load measured by physiological measures, to the best of our knowledge no studies about the cognitive load during a game, measured by physiological measures have been performed.

1.3 Psychophysiological measures

There are many different physiological measures that are used in psychological research; a few examples are electrodermal activity (EDA), heart rate (HR), and heart rate variability (HRV). These measures can be used because physiological changes occur when a human is in danger and a stress reaction is created. This danger ensures that the body can adapt to the stressor and that you have more energy to fight or flight. In these modern times these reactions also occur when people experience psychological stress or when playing video games (Campbell & Ehlert, 2012; Hébert, Béland, Dionne-Fournelle, Crête, & Lupien, 2005). Physiological changes occur also naturally, for example with temperature changes. When people experience psychological stress, their subjective feelings about that stress level correlate with their physiological responses. According to Dishman et al. (2000), there is an inverse relationship between HRV and perceived emotional stress. This would mean that when the perceived stress increases, the stress level rises, and the HRV becomes smaller. Subjective, perceived stress could thus be related to the stress reactions in the body (Campbell & Ehlert, 2012).

There are two components that play a part in a stress reaction, the hypothalamic-pituitary-adrenal (HPA) axis and the sympathetic nervous system, which are (together with the parasympathetic system) part of the autonomic nervous system. The sympathetic and parasympathetic system work at the same time but in opposite directions (Setz et al., 2010). This means that when a stress reaction occurs, the sympathetic nervous system activates physiological responses, and the parasympathetic nervous system inhibits those responses. The reason for this is that the sympathetic nervous system is considered the ‘fight or flight’ system, and the parasympathetic nervous system the ‘rest and digest’ system. When your body is in danger, it activates the muscles and heart rate, but digesting your last meal is something that can wait.

Every organ in our body is linked to both systems, except the skin. The sweat glands and skin blood vessels are exclusively innervated by the sympathetic nervous system. This means that when a (psychological) stress reaction occurs, the skin gives a response and it is therefore an ideal measure for changes in arousal (Setz, Arnrich, Schumm, Marca & Tröster, 2010). The heart and the skin are linked to both systems (Setz et al. 2010). In many studies this has led to the result that skin conductance and heart rate both increased when subjects were exposed to arousing stimuli (Loggia, Juneau, & Bushnell, 2011). Jacobs et al. (1994) found that, as with indexes of cardiac performance such as HR, EDA reliably increases during laboratory mental stress and under threatening, novel or challenging conditions. These findings, together with the fact that skin conductance, HR, and HRV are innervated by the sympathetic nervous system, led to the expectation that there could be a relation between EDA, HR, and HRV.

1.3.1 EDA

EDA is a direct measure of sympathetic activity; and is less ambiguous than facial muscle and heart activity (Kivikangas et al., 2011). It refers to the electrical properties of the skin, which vary in response to sweat secretion by the sweat glands. When using a low constant voltage, changes in skin conductance can be measured. When the skin is sweatier the conductance will be higher. Measuring this is mostly done at the palmar sites of the hands or feet because the density of eccrine sweat glands is highest in this area of the body (Setz et al., 2010; van Dooren, de Vries, & Janssen, 2012). EDA signals are divided into three different types.

The first type is slow changing signals, called the *skin conductance level* (SCL). This measures the psychophysiological activation and gradual long-term shifts in conductance. These signals can vary substantially between individuals and are therefore not useful when comparing many participants (Kappeler-Setz, Gravenhorst, Schumm, Arnrich, & Tröster, 2011). The SCL seems to be related to the level of alertness or attention and task demand (Miller, 1979).

The second type is a fast changing peak; these peaks occur in reaction to a single stimulus and are called (specific) *skin conductance responses* (SCR). These peaks occur 1 - 5 seconds after the stimulus. To describe the SCR, the amplitude of the SCR, the latency (between stimulus and SCR onset), and the recovery time are used. The SCR signal is a good and pure measure for emotional arousal (Miller, 1979). However, SCRs are slow (they have a delay from one to five seconds), thus it is important to know which stimuli signals belong to.

Besides the SCR, there are also *non-specific fluctuations* (NS.SCR) in the EDA data. These fluctuations occur in the absence of an external stimulus (Setz et al., 2010), and the typical measure of these fluctuations is the frequency per minute. A typical value is 1 – 3 NS.SCR per minute in rest (Dawson, Schell, & Filion, 2007). The frequency in rest can be measured by using a baseline before an experiment starts.

1.3.2 HR and HRV

The heart is part of the cardiovascular system and regulates the blood flow in the body (Drachen, Pedersen, Nacke, & Yannakakis, 2010). Darwin (1886) stated:

> "...when the mind is strongly excited, we might expect that it would instantly affect in a direct manner the heart…when the heart is affected it reacts on the brain; and the state of brain again reacts through the pneumo-gastric nerve on the heart; so that under any excitement there will be much mutual action and reaction between these, the two most important organs of the body" (p. 69).

This implies that many years ago it was already expected that the heart and the brain were related to each other, and that when a person was psychologically aroused, the heart reacts to that arousal. To get insight in the reactions of the heart, HR is commonly used. HR is the number of heart beats per minute and is often used as a physiological measure of physical workload, arousal or emotions (Drachen et al., 2010).

HRV is a measure of the continuous interplay between the sympathetic and parasympathetic influences on HR. It gives us information about autonomic flexibility, and thus about the capacity for regulated emotional responding (generating emotional responses of appropriate timing and magnitude) (Appelhans & Luecken, 2006). Ravaja, Saari, Salminen, Laarni and Kallinen (2006) showed in different studies that tasks that require cognitive effort provoke arousal and activate the sympathetic nervous system. This is pointed out by a decrease in HRV. Stress has the same influence on HRV. When subjects are stressed the HRV is suppressed, but when subjects are relaxed, HRV emerges (Mandryk, Inkpen, & Calvert, 2006). HRV consists of different measures that calculate the variability between heartbeats. When measuring HRV it is necessary to detect each heartbeat. The time between heartbeats is called the interbeat-interval (IBI). In figure 1 (Appelhans & Luecken, 2006, p.232) the IBI is shown.

HRV can be measured by using different kinds of measures like time domain measures and frequency domain measures. There are also different time domain measures like the standard deviation of the interbeat-interval (SDNN), the root-mean-square successive difference (rMSSD) that calculates the square root of the mean of the squared differences between successive interbeat-intervals, and the pNN50 that calculates the percentage of differences between successive interbeat-intervals that are greater than 50ms (Bilchick & Berger, 2006). In this study, the SDNN will be used because the rMSSD, pNN50 and frequency domain measures were beyond the scope of this thesis.

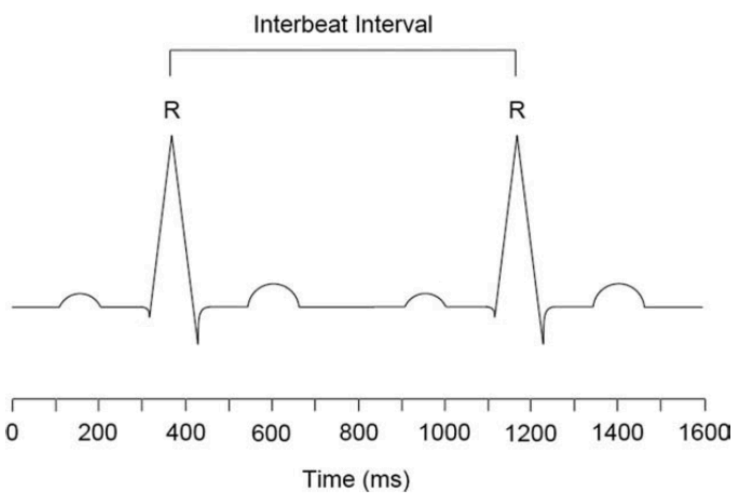

Figure 1. Idealized electrocardiograph segment representing two heartbeats. Waveforms are labelled with letters and correspond with specific electrophysiological events during a heartbeat. The interbeat-interval is defined by the temporal distance between R-spikes, the waveforms corresponding to depolarization of the heart’s ventricles. Reprinted from “Heart Rate Variability as an Index of Regulated Emotional Responding”, by B. M. Appelhans, and L. J. Luecken, 2006, *Review of General Psychology*, 10 (3), 229-240. Copyright (2006) by the American Psychological Association.

1.4 Gaming and psychophysiological measures

Little is known about psychophysiological responses to serious games, although there have been studies about psychophysiological measures in combination with ‘normal’ games such as first-person shooter games. In a study where they used first-person shooter games to measure physiological responses, Drachen et al. (2010) discovered that there is a correlation between psychophysiological measures and gameplay experience and that games can evoke physiological responses. Carnagey, Anderson, and Bushman (2007), who studied the effects of playing a violent video game on physiological measures, demonstrated that playing a violent video game for 20 minutes could cause people to become less physiologically aroused by real violence. There are also studies that investigated physiological measures by non-violent games. For example, Cusveller, Gerritsen, and de Man (2014) who studied the effects of three different, simple, games on skin conductance and heart rate. They increased the pace of the games until the player was ‘game over’ and increases in skin conductance were found. It has also been shown that when playing a video game (especially a violent video game), heart rate increases. However it should be kept in mind that the heart is not only innervated by the parasympathetic nervous system, but also by the sympathetic nervous system. This can make interpreting the results difficult because increased cardiac sympathetic activity is related to emotional arousal, which increases the heart rate. On the other hand increased parasympathetic activity is related to information intake and attentional engagement. This activity decreases the heart rate (Ravaja et al., 2006).

Investigating the effects of games on psychophysiological data can be difficult as physiological processes are never related to psychological phenomena in a one-to-one relationship. This means one can easily over-interpret the meaning of physiological signals. Games are not the simplest stimuli that can be used. They have different kinds of output (visual and auditory) and sometimes require the use of physically complicated input devices (like a joystick). Different complicated cognitive processes are also required that operate on times scales from seconds to several hours (Kivikangas et al., 2011). Another problem is that there is no general agreement about how game experience arises, what could be used in psychophysiological game research. Thus research questions have to be composed that can be answered even when the game is complex or when there are many psychological processes.

1.5 Purpose of this study

The external stakeholder for the Master thesis is T-Xchange, a joint venture between Thales Netherlands and Twente University. They are interested in gaining insight into serious game artefacts and gaming that facilitates human judgment and decision making regarding social challenges. They want to understand how complex social systems evolve over time, how to develop game artefacts that through interactions create player experiences, and how to transfer these experiences to real world situations (“T-Xchange,” n.d.). As a result of this, they are curious about the physiological reactions of players who are playing a serious game. They are interested in what the effects of game experiences are on the level of arousal and if these effects can be influenced. They hope that it is possible to intervene when they see, real-time, that a player experiences too much or too little arousal so the games can adapt to different players.

The main goal of this study is to investigate if it is useful to monitor physiological measures in real-time. Therefore we made two adapted versions of a serious dilemma game of T-Xchange, a slow version and very fast version that includes the sound of a ticking clock and has a shorter playtime. We want to see if the physiological reactions are different between those two versions. If that is the case than future research can be done about adapting the game to real-time physiological arousal.

The first research question that we want to answer is: **Does the number of NS.SCR differ between the two versions of the game?** In this research the focus is on NS.SCR because we look at the whole game instead of at specific events in the game. We will compare the differences in NS.SCR per 5 minutes between two different gaming situations, a slow version and a fast version. Hébert, Béland, Dionne-Fournelle, Crête, and Lupien (2005) studied the effect of sound on physiological activity in video games. They discovered evidence that players who played a video game with sound had higher cortisol levels, which was assumed to reflect the stress induced by the game. Therefore it is expected that the fast version that includes the sound of a ticking clock will give more stress, what can result in an increased amount of NS.SCR.

By decreasing the playtime of the fast version and by making dilemmas appear faster, the task difficulty can be increased. This led to a secondary task; because the participants had to focus on a dilemma while other dilemmas popped up simultaneously, the participants had to divide their attention. Adding sounds and increasing task difficulty also increased the cognitive workload. It is expected that this would lead to more NS.SCR per 5 minutes in the faster version of the game (Ang et al., 2007). Cusveller, Gerritsen, and de Man (2014) found similar results. When they increased the level of game difficulty, skin conductance increased too. Because our faster version of the game can be seen as more difficult, we expect to see more NS.SCR in that version.

The second research question that we want to answer is: **Do the HR and HRV differ between the two versions of the game?** Mehler, Reimer, Coughlin, and Dusek, (2010) reinforced older studies where it was discovered that HR is a sensitive measure for cognitive workload in a driving related research. The effectiveness of HR was also shown in research in the field of gaming (Carnagey et al., 2007). Because of this effectiveness it is expected that there will be differences in HR between the two gaming situations. In the faster version the cognitive workload will be higher because of the increased task difficulty and it is expected that the HR will be higher in that version of the game compared to the slow version. Furthermore, HRV will be used because Derbali and Frasson (2010) studied the motivation of players who were playing a serious game and they suggested that for future work adding more variables like HRV could improve their models predictions. It is expected that HRV will be smaller in the fast version, because when people are stressed or more aroused, HRV is suppressed (Mandryk et al., 2006).

The third research question that we want to answer is: **Is there a relation between the HR, HRV, and EDA during the game?** A (negative) correlation is expected between the HRV measure (SDNN), and the NS.SCR per 5 minutes. It is expected that the SDNN will be shorter when the number of NS.SCR increase. This was discovered by Loggia, Juneau, and Bushnell (2011), who found that graded intensities of painful heat stimuli evoked increases in both heart rate and skin conductance. We expect that when there is a relation between physical stress and physiological measures, cognitive load can provoke the same response (Jacobs et al., 1994).

The last research question that we want to answer is: **Is there a relation between the physiological measures and subjective stress?** Dishman et al. (2000) found an inverse relationship between HRV and perceived emotional stress. This would mean that when the perceived stress increases, the HRV decreases. This research raised the question whether our experiment could also show a relation between HRV and perceived stress. We also expected a relation between HRV, HR, and EDA, and therefore we would like to investigate if a relation can be found between our physiological measures and the subjective stress during the game.

2. Method

2.1 Participants

A total of 43 participants took part in this experiment of which 18 were male and 25 female. Their age ranged between 18 and 30 with an average of 21 (SD = 2.8). 25 participants were Dutch, 17 German and 1 of which the nationality was unknown. All of the participants were students and most were first years (53.5%), followed by students who were in their first year of their master (20.9%). The other participants were spread over the other study years. The ethical committee of the University of Twente approved this study. Before participants took part in this research they filled in an Informed Consent form that is included in appendix A.

2.2 Materials

2.2.1 Empatica E3™ sensor

To record EDA and HRV the Empatica E3 sensor was used. Figure 2 shows the Empatica E3 sensor attached to the wrist of the non-dominant hand. The E3 (46 x 40.8 x 13.8 mm; weight 25 grams) is a small device that consists of a little black box on top of the wrist that measures heart rate (by using a green bright light to shine through the skin) and an EDA sensor at the other side. The band can be closed with Velcro. The E3 measures at a sampling rate of 4Hz. Electrodes were used that do not need additional electrode paste. After each gaming session, the EDA and IBI data was transmitted to a personal computer by using the Empatica Connect software. This software uploaded, securely encrypted, the data to the Empatica website where the data could be downloaded as an Excel file.

Figure 2. The Empatica E3 on the wrist of a non-dominant hand.

2.2.2 The mayor dilemma game

The Mayors game (see figure 3a) is an SDM (strategic decision making) type of game and is a web-based game used in training sessions. These training sessions are developed for gaining experience, stimulating reflection and awareness and sharing experiences.

Figure 3a. A screenshot of the mayor dilemma game on which the student dilemma game was based. In this screenshot the first dilemma is shown, with two more dilemmas waiting to be opened. For more information the player can click on the three advisors with an white ‘i’ above them.

Based on the Mayors game, there was a new SDM scenario created for students so they could be used as participants in this research (see figure 3b). A scenario was created that takes into account the time management problems that students often experience during the semesters and during preparation for exams. This scenario has the same structure as the mayor dilemma game and consists of six different dilemmas about distractions, problems with your study planning and choosing between fun things to do and doing what is best for your study. The player can choose which dilemma will be answered first, but the dilemmas will be available in a fixed order. The game ends when the time is up or when all the dilemmas have been answered. In these dilemmas there are six advisors sitting at a table, which give information and advice about their opinion about the dilemma. A friend, teacher, study adviser, mother, chairman of your sports club, and your conscience were chosen as advisors. Each person will give his or her opinion once or twice. After the dilemma has been answered the player has to fill in which opinion influenced him or her in making the decision.

Figure 3b. A screenshot of the student dilemma game. The structure of this scenario is the same as the ‘Mayors game’ shown in figure 3a. Only the theme, figures, dilemmas, and information are different.

The experiment has two different versions, one slow and one fast version. The slow version has a low information density. The time between the dilemmas is relatively long (round 2 minutes), the time between the information items that the advisors give is longer (round 30 seconds), and the players get twenty minutes to complete all the dilemmas.

The fast version has a high information density. The time between the dilemmas is average 1 minute shorter, the people around the table give their information quickly after each other, sometimes at the same time, and some characters give information twice in one dilemma. Besides that, the players get six minutes to complete all the dilemmas, and there is a sound of a ticking clock that should increase the feeling of time pressure and thus is expected to give more arousal. A description of the scenario can be found in Appendix B.

2.2.3 Survey

Besides the physiological measures we also wanted to know how the participants experienced the game. Therefore a survey (see appendix C) with demographic questions, questions about the experienced stress, questions about the game, and general questions about stress was taken by the participants when the experiment was completed. The survey consisted of closed questions, questions with a Likert scale from 1 to 10, or mulitple choice questions.

2.3 Procedure

Two different game versions were made (see section 2.2.2). The fast version should lead to a higher arousal (higher skin responses) compared to the slow version. For both versions the procedure was the same.

The participant came to the research room where the instructions about the experiment were explained. To get a better understanding of the game, a demo-game was played for a few minutes. When it was clear what the purpose of the experiment was and how the game worked, the E3 was placed on the non-dominant wrist and the participant took place behind the personal computer. The informed consent form was signed and the instructions of the game were explained. The E3 was set to recording mode and the game was started.

Two baselines were included, one at the start and one at the end of the game. The baseline consisted of calming music that was played for 5 minutes (a link to the music video can be found in appendix D). When the music ended the dilemmas started automatically. When a dilemma was finished there was a screen that asked on a scale of 3 how stressed the participant was. When this question was completed the next dilemma appeared. When all dilemmas were answered (this period lasts around 7 minutes in the fast scenario and around 10 in the slow scenario), the second baseline began. The same calming music was played again for 5 minutes. After the second baseline the scenario ended automatically.

As the E3 records timestamps and the game also logged timestamps of game events and actions, the data from the sensors and the game could be synchronized. The logfiles of the game were used to know the exact time of the start of the two baselines and dilemmas.

When the game ended the participant had to fill in a survey about the subjective stress levels during the game and about daily stress.

2.4 Data analysis

Different analyses were used in this study. The software that was used for the statistical analyses was IBM SPSS 22, and the significance level was set to $\alpha = .05$. For analysing the HR and HRV data the free Kubios HRV software was used (Tarvainen, Niskanen, Lipponen, Ranta-Aho, & Karjalainen, 2014). In this software the time-domain method was used to calculate the mean HR, the interbeat-interval, the standard deviation from the HR, and the standard deviation from the interbeat-interval (SDNN).

To analyse the data for the first three research questions the outliers and errors were deleted. When using the NS.SCR variable, participants were removed when more than two of the three game parts (first baseline, game, second baseline) contained no NS.SCR. Instead of 43 participants, 32 participants were analysed. When using the variables HR and HRV, the data from the participants that showed measuring errors were removed. This was done by analysing the number of measurements in each dataset. The five datasets with the smallest amount of measurement points were removed because these contained far fewer measurements points than the other datasets. Examples of a dataset that lacked in the amount of measurement points, and an example of a correct dataset can be found in figure 4 and 5 in appendix E. As a consequence the sample size was reduced to 38. When these variables were combined, for testing the correlation between HRV and EDA, these deselected subject numbers were also combined. The subjects were reduced to 28.

To determine if the number of NS.SCR differ between the two versions, and the different parts of the game, a Repeated measures Two-Way ANOVA was performed. The variables that were used were the number of NS.SCR during the game, the number of NS.SCR in the first baseline, and the variable 'version', that consisted of the fast and slow version. Before this test was performed the assumption was checked whether the data was normally distributed. This requirement was fulfilled (see figure 6, appendix F).

To investigate if the HR and HRV also differ between the different versions and parts of the game a Repeated measures Two-Way ANOVA was performed. The variables that were used were the mean HR in the first baseline, the mean HR during the game, the mean SDNN in the first baseline, the mean SDNN during the game, and the variable 'version'. Before this test was performed the assumption was checked whether the data was normally distributed. This requirement was fulfilled (see figure 7 and 8, appendix F).

The first baseline was exactly the same for the fast and slow scenario, so no differences in HR were expected. We checked this with an independent samples t-test, because the dependent variable 'Mean HR during the first baseline' was approximately normally distributed (see figure 9, appendix F).

To check if there is a correlation between HR, HRV, and EDA changes, Linear Regression was used.

To answer the question if the physiological measures and subjective stress levels correlate with each other, we first looked at the surveys to see if the participants who played the fast version experienced more stress during the game. First the assumption if the dependent variable, the experienced stress during the game, is normally distributed was checked (see figure 10, appendix F). The data was not normally distributed so the Mann-Whitney U test was performed. To check the correlation between the physiological measures and the subjective data Spearman's Rho correlation coefficient was used because the subjective stress during the game is an ordinal variable.

3. Results

In this chapter the results for the four different sub-questions are shown.

3.1 Does the number of NS.SCR differ between the different versions and parts of the game?

Although the number of NS.SCR in the fast version (M = 5.5, SD = 3.3) was higher than in the slow version (M = 5.2, SD = 3.3), the main effect for the different versions of the game was not significant; F (1, 31) = 1.4, p = .2. There was also a difference between the different parts of the experiment. During the game (M = 5.3, SD = 3.0) there were more NS.SCR per 5 minutes than during the first baseline (M = 4.5, SD = 3.2) or during the second baseline (M = 4.7, SD = 3.0). However, the main effect for the three different parts of the experiment was also not significant; F (2, 29) = <1. This means that the number of NS.SCR was not affected by the different versions of the game of the different parts of the experiment. The different versions and the different parts of the experiment could interact with each other; therefore it was investigated if there was an interaction effect. There was no interaction found between the different parts and versions; F (2, 96) = <1.

3.2 Do the HR and HRV differ between the different versions and parts of the game?

When looking at the HR between the fast and slow version, the HR in the slow version (M = 80.6, SD = 10.3) was significantly higher than the HR in the fast version (M = 76.1, SD = 7.7); F (1, 37) = 4.5, p = .0. Thus, there was a main effect for the variable 'version' on HR. This means that the HR differs between the two versions of the game. The HRV was less (M = 65.2, SD = 19.4) in the slow version than in the fast version (M = 68.4, SD = 17.8). However, this difference was not significant; F (1, 37) = <1, so there was no main effect for the variable 'version' on HRV, what means that the HRV did not differ between the two versions of the game.

It was found that the HR in the first baseline was higher (M = 79.6, SD = 10.3) than the HR in the game (M = 78.3, SD = 9.3). This difference was significant; F (2, 36) = 4.4, p = .04. So there was a main effect for the different parts on HR. The HRV in the game was smaller (M = 66.8, SD = 18.4) than in the first baseline (M = 69.3, SD = 21.4), but there was no main effect found for the variable 'part'; F (2, 36) = <1. This means that the HRV was the same in the baseline as in the game.

When looking at the interaction between the variables 'version' and 'part', there was a interaction effect for HR; F (2, 1332) = 9, p = .005. However, there was no interaction effect for HRV; F (2,1332) = <1. This means that the HR in the baseline increased even more when a participant played the slow version of the game.

Lastly, during the first baseline, the HR in the fast version (M = 75.5, SD = 7.4) was lower than in the slow version (M = 83.6, SD = 11.4). This difference was tested and it can be concluded that there is indeed a significant difference in HR during the first baseline; t (36) = -2.6, p = .02.

Figure 11.

This figure shows the difference in HR between the baseline and the game, for the slow and fast version, with the standard error of the estimated means with 95% CI intervals. In both parts of the experiment the HR is higher in the slow version than in the fast version. The fast version shows an expected pattern; during the game the HR was higher than in the first baseline. However, in the slow version an unexpected effect was seen; the HR was higher in the baseline than in the game.

3.3 Is there a relation between the EDA, HR, and HRV during the game?

There was no correlation between EDA and HR; r = -0.1, p = .7. There was also no correlation between EDA and HRV; r = 0.1, p = .5. This means that when the number of SCR increased, the HR did not increase and the HRV did not decrease. Detailed graphs of the relations between HR, HRV, and EDA, can be seen in appendix G.

3.4 Is there a relation between the physiological measures and subjective stress?

When looking at the question 'how much stress did you experience during the game?' we see that the players of the fast scenario (M = 4.2, SD = 2.0) experienced more stress than the players of the slow scenario (M = 3.3, SD = 1.9). However, this difference in experienced stress was not significant; $U = 168$, $p = .1$.

When comparing EDA, HR, and HRV to the experienced stress during the game, there was no correlation between the experienced stress and the number of NS.SCR; $\rho = -.1$, $p = .6$. There were also no correlations found between the experienced stress and the HR; $\rho = -.3$, $p = .1$, and the HRV; $\rho = .1$, $p = .5$. This means that when the experienced stress increased, the EDA and HR did not increase too, and that the HRV did not decrease. Scatterplots that show the relation of the data between the physiological measures and the experienced stress can be seen in appendix H.

4. Discussion

The main goal of this study was to investigate if it is useful to monitor physiological measures during a serious game in real-time.

4.1 Different versions, different effects?

Our first research question was 'Does the number of NS.SCR differ between the two versions of the game?' We can conclude that we did not find any statistical differences in the number of NS.SCR between the two different versions of our game. This was not in line with our hypothesis, namely that there would be more NS.SCR in the fast version. However, this was found in the study of Cusveller et al. (2014) where they measured differences in skin conductance by games that increased in difficulty. Their games increased in difficulty by going faster and switching faster between different games. Therefore it was thought that our faster game would lead to an increase in NS.SCR as well. Based on the study of Hébert, Béland, Dionne-Fournelle, Crête, and Lupien (2005), it was expected that the sound of the ticking clock in the fast version (and a higher speed of stimulus presentations), would lead to more arousal. Although there were slightly more NS.SCR in the fast version, this difference was not significant. Comparing this experiment to the study of Cusveller et al. (2014), we can conclude that they increased the cognitive workload of their participants more than we did, as their games increased in speed until the player was 'game over' and their games were more difficult to comprehend. In their research the skin conductance increased with each game that was harder to play. According to Mehler, Reimer, Coughlin, and Dusek (2010) an increase in skin conductance level is a good indication of the change in workload. As we did not find any differences in EDA, this could be an indication of no change in cognitive workload.

Looking at the different parts of the experiment, no difference was found between the two baselines at the begin and end of the experiment, and the game. This could be due to the fact that the game did not evoke enough arousal. In the study of Cusveller et al. (2014), they used more entertaining games instead of serious games like 'Guitar Maniac' where players had to play guitar as fast as possible. This game had a scoring system, and counted the number of mistakes, which could have led to more arousal than our game.

The second research question was 'Do the HR and HRV differ between the two versions of the game?' It can be concluded that there was a significant difference between the two versions. The HR in the slow version was higher than in the fast version. However, when we compare the HR of 80 in the slow version to the average HR of an adult (between 60 – 70 beats per minute when resting), this is not a large increase. It is peculiar that the slow version of the game evokes this relatively higher HR of 80, though it could be the case that this version was too slow and that it caused boredom. According to London, Schubert, and Washburn (1972) HR (and skin conductance) can increase when someone is bored, which can lead to stimulus seeking (Berlyne, 1960). Another possibility for this result is that the slow version caused frustration in the players. According to Mandryk and Atkins (2007) increased HR can be related to arousal and increased arousal can be related to frustration. This makes sense, as in our research the participants who played the slow version often made comments like: 'does it still work?' or 'is the game already done?' Some of them began to move like they were in a hurry and could not wait for the experiment to be over. This implies that they might have been frustrated, which could have caused the increased HR in the slow version of the game.

Looking at the different parts of the experiment, we found differences between the baseline and stressful part of the experiment. During the first baseline the HR was significantly higher than during the game. Compared to Kirschbaum et al. (1993), the difference that we found was the other way around. In their research they found more physiological arousal during the stressful part of the experiment. This result in our study was unexpected and could be an indication that there was more workload during the baseline than during the game. However, when we look closer, this difference in HR was not that big and it fits in the HR range of a normal adult. HR is therefore not a good indication of a higher cognitive workload during the baseline. Although the difference fitted in the HR range of an adult, a decrease in HR during the game could also be an indication for increased parasympathetic activity. According to Ravaja et al. (2006), this is related to information intake and paying attention. Thus the decrease in HR that we found during the game could be due to the fact that the participants were alert. Concluding, we can say that the game could have evoked frustration (in the slow version) and alertness (in the fast version). These states could have led to the results that we found.

An unexpected result was that in the first baseline of the experiment, the mean HR in the slow version of the game was significantly higher than in the fast version. We did not expect to find this result because the baselines are exactly the same in each version. A possible explanation could be that this difference is due to coincidence. It could be that some participants arrived stressed at the experiment (from biking, or were in a hurry) and happened to play the slow version of the game, whereas in the fast version of the game there were no participants that were already stressed before the experiment. In that case, the randomization of the participants failed and this could have influenced the results in our study.

Looking at the HRV, we found a smaller HRV in the slow version but this effect was not statistically significant. This can be an indication that the participants felt more stressed or aroused in the slow version of the game, as according to Mandryk et al. (2006) HRV is suppressed when people are stressed. In combination with the significant HR results, the results that we see in the HRV could be a sign of frustration due to the fact that the slow version was too slow.

4.2 Relation between EDA, HR, and HRV

Our third research question was 'Is there a relation between the HR, HRV, and EDA during the game?' Although a positive relation between NS.SCR and HR and a negative correlation between SCR and HRV were expected, no relations between these measures were found. To a great extent this was due to the fact that there were very few NS.SCR. According to Dawson et al. (2007) a typical value is 1 – 3 NS.SCR per minute in rest. Compared to our study, where approximately 1 NS.SCR/min were found (during the baseline and during the game), this number is low. This lack of NS.SCR could have influenced the relation between those measures negatively, as it is unclear if this is the result of the sensor not working well, or the game not evoking enough stress reactions.

Loggia, Juneau, and Bushnell (2011), did find some relations between heart rate and skin conductance. However, their research had more intense stimuli than ours. They discovered that graded intensities of painful heat stimuli evoked increases in both heart rate and skin conductance. It was expected that same results could be found because the skin and heart are both innervated by the sympathetic nervous system. However, it is possible that our stimulus, the game, did not cause changes that were noticeable enough because mental stress can be less arousing than physical stress.

4.3 Self report compared to physiological measurements

For the fourth research question we compared the change in subjective stress to the change in physiological measures (number of NS.SCR, HR, and SDNN). There were no correlations found between the subjective and objective measurements of stress, even though these correlations were expected according to Dishman et al. (2000) and Campbell and Ehlert (2012). Dishman et al. (2000) found an inverse relationship between HRV and perceived emotional stress. This would mean that when the perceived stress increases, the HRV decreases. Although earlier studies suggested that significant correlations occasionally occur, there is a general agreement that physiological measures and subjective report do not correlate well (McLeod, Hoehn-Saric, & Stefan, 1986). McLeod et al. (1986) found that subjects could determine the direction of the change. When the subjective stress increased, the physiological measures increased too, however they could not indicate the degree of increase correctly. A recent study confirms this agreement (Cusveller et al., 2014). They studied the effects of arousal in games and also found no correlation between subjective stress and physiological measures. Our study is most similar to that of Cusveller et al. (2014), therefore our results are not unusual. It is possible that when people are playing a video game, they are not aware of the physiological reactions in their bodies and what the effects of games are on their stress level.

4.4 Limitations

In this research there were some limitations that could have influenced the results. The first limitation is that the E3 sensor from Empatica did not work very well. Many participants had to be excluded from the study because the EDA, HR and HRV were often badly measured (a comparison between a bad and good dataset can be found in appendix E). Besides the sensors not working as expected, the ergonomics of the product did not meet our expectations either. The heart rate sensor on top of the wrist was not formed in the shape of a wrist. Therefore the heart rate sensor did not fit properly on most of the participants. Another influencing factor could be the wristband of the E3, as this wristband was made of Velcro and could be ordered in different sizes. An M/L size was ordered and cut into a size that was thought to fit most wrists. The skin of the inner side of the wrist was inspected after the experiment to check that the EDA sensors had made good contact (seen by the imprints of the sensors in the skin). This was the case for each participant. However, when the participant had a skinny wrist, the upper sensor for the heart rate floated slightly above the skin. This could have influenced the results in a negative way and led to a decrease in the number of valid datasets. These limitations of the E3 sensor were not known before it was ordered. It was thought that this sensor would be better than the older, discontinued, Affectiva Qsensor because it combines measuring EDA, HR, and HRV.

The second limitation of this study was the type of serious game. The dilemma game was not one of the most exciting games that could be played. Compared to Drachen et al., (2010) and Ravaja et al., (2008) where they used more violent games, this game did not have many actions in it. Looking at the study of Cusveller et al. (2014), it is clear that the adaptations that they made to their games were more stressful than our adaptations. We can learn from their study that for significant EDA responses, a game has to be faster and has to have more secondary tasks.

The last limitation was that the baseline in this experiment did not work well. There was a difference in HR between the first baseline and the game, whereby the HR in the baseline was higher. Besides that, there was also a difference in HR in the first baseline, between the fast and slow version of the game. This is an indication that our baseline was not relaxing enough. The fact that the participants were not totally at rest before the experiment started could have led to an incorrect skin conductance level and a higher HR in rest. The number of NS.SCR was not influenced because even when the SCL is higher, the responses are still visible.

5. Recommendations

For future research the older Affectiva Qsensors are recommended. If these old sensor are unavailable it is recommended to test the new E4 from Empatica. This new sensor was released a few months after this experiment and is built more ergonomically and is usable for different wrist sizes. It is expected that when these sensors are used that the problem with the wrong wrist sizes will disappear. This would reduce the number of invalid datasets.

The second recommendation is to study in further research what the exact point is where participants are optimally aroused. A model that can be interesting to take into account in this future work is the Yerkes-Dodson law, also known as the inverted-U-shape function (Salehi, Cordero, & Sandi, 2010). This model shows that when arousal increases, the performance increases too, up to a specific point. This point is called the ‘optimum level of arousal’. It would be interesting to know when someone reaches the optimum level of arousal, because that is the best condition to learn (Salehi, Cordero, & Sandi, 2010). This result is consistent with Joëls, Pu, Wiegert, Oitzl, and Krugers (2006), who discovered that stress affects learning and memory processes. According to them, focused attention is induced when stress is experienced within the context of a learning experience. This can indicate that when players are feeling stressed that they are more focussed than when they would are relaxed. The latter can be very interesting for a compagny like T-Xchange. When the players of their games are feeling exactly the right amount of stress, their game performance can increase and the learnability of the serious games will probably increase. This can be studied by using a more arousing game. T-Xchange does have some serious games where short movies are shown between the dilemmas. These movies contain footage of disasters and news items. By making the player feel like he or she is participating in a realistic experience (called immersion), the transfer of learned information can increase (Dede, 2009). This means that players will learn more from the trainings and that they can use those learned information better in real situations.

The last recommendation is to make a high-stakes serious game. This means that there are more important consequences for the player. The players will feel more pressure and this can lead to more physiological arousal. However, this can change the purpose of the game and a consequence of this is that the players become more competetive, which could negatively affect the learnability.

6. Conclusion

The main goal of this study was to investigate if it is useful to monitor physiological measures during a serious game in real-time. We can conclude it is very hard to get reliable results from real-time physiological data when players are playing a serious game. For this reason, we do not recommend adapting a serious game based on physiological data.

We expected to see more arousal in the fast version of the game. Even though we did find differences in HR, we did not find this in the skin conductance. What is surprising is that the difference in HR was the other way around than expected: HR increased in the slow version of the game. This could have been due to frustration, but studies also show that HR can decrease when participants are paying attention. Keeping in mind that the Empatica E3 sensor did not work very well, it is hard to say which factor caused this difference.

For further studies it is recommended to investigate the point where participants are optimally aroused so that player performance can be increased. This can be studied by using a serious game of T-Xchange that already includes video montages of disasters, which can increase the transfer of learned information. It is also recommended to make a more high-stakes game that has more consequences for the player. However, this can change the purpose of the game negatively.

Overall, when we look at what is happening in the world, we can see that measuring real-time information of the body is becoming increasingly popular. More and more devices that measure HR, skin conductance, and your sleep patterns are emerging on the market for individual use (for example the Jawbone UP3 and Apple Watch). Our research shows that when someone is measuring his or her health during the day, that these results have to be interpreted with caution.

7. References

Ang, C. S., Zaphiris, P., & Mahmood, S. (2007). A model of cognitive loads in massively multiplayer online role playing games. *Interacting with Computers*, *19*, 167–179. doi:10.1016/j.intcom.2006.08.006

Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology*, *10*(3), 229–240. doi:10.1037/1089-2680.10.3.229

Berlyne, D. E. (1960). Toward a Theory of Exploratory Behavior: 1. Arousal and Drive. In *Conflict, Arousal, and Curiosity* (pp. 163–192). New York: McGraw-Hill Book Compagny.

Bilchick, K. C., & Berger, R. D. (2006). Heart Rate Variability. *Journal of Cardiovascular Electrophysiology*, *17*(6), 691– 694. doi:10.1111/j.1540-8167.2006.00501.x

Brookhuis, K. A., & De Waard, D. (2001). Assessment of drivers' workload: Performance and subjective and physiological indexes. *Stress, Workload, and Fatigue*, 321–333.

Campbell, J., & Ehlert, U. (2012). Acute psychosocial stress: does the emotional stress response correspond with physiological responses? *Psychoneuroendocrinology*, *37*(8), 1111–34. doi:10.1016/j.psyneuen.2011.12.010

Carnagey, N. L., Anderson, C. A., & Bushman, B. J. (2007). The effect of video game violence on physiological desensitization to real-life violence. *Journal of Experimental Social Psychology*, *43*(3), 489–496. doi:10.1016/j.jesp.2006.05.003

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, *59*(2), 661–686. doi:10.1016/j.compedu.2012.03.004

Cusveller, J., Gerritsen, C., & de Man, J. (2014). Evoking and Measuring Arousal in Game Settings. In *Games for Training, Education, Health and Sports* (pp. 165–174). Springer International Publishing.

Darwin, C. (1886). *The expression of the emotions in man and animals (3rd ed.)*. (D. Appleton, Ed.)*The expression of the emotions in man and animals*. New York. doi:10.1037/10001-000

Dawson, M., Schell, A., & Filion, D. (2007). The Electrodermal System. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (pp. 159–181). Cambridge Univeristy Press.

Dede, C. (2009). Immersive Interfaces for Engagement and Learning. *Science*, *323*(January), 66–69.

Derbali, L., & Frasson, C. (2010). Prediction of Players Motivational States Using Electrophysiological Measures during Serious Game Play. *2010 10th IEEE International Conference on Advanced Learning Technologies*, 498–502. doi:10.1109/ICALT.2010.143

Dishman, R. K., Nakamura, Y., Garcia, M. E., Thompson, R. W., Dunn, A. L., & Blair, S. N. (2000). Heart rate variability, trait anxiety, and percieved stress among physically fit men and women. *International Journal of Psychophysiology*, *37*, 121–133.

Drachen, A., Pedersen, A. L., Nacke, L. E., & Yannakakis, G. (2010). Correlation between Heart Rate , Electrodermal Activity and Player Experience in First-Person Shooter Games. In *Sandbox '10 Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games* (pp. 49–54).

Hébert, S., Béland, R., Dionne-Fournelle, O., Crête, M., & Lupien, S. J. (2005). Physiological stress response to video-game playing: the contribution of built-in music. *Life Sciences*, *76*(20), 2371–80. doi:10.1016/j.lfs.2004.11.011

Ikehara, C. S., & Crosby, M. E. (2005). Assessing Cognitive Load with Physiological Sensors. In *HICSS'05. Proceedings of the 38th Annual Hawaii International Conference on System Science.* (Vol. 00, pp. 1–9).

Jacobs, S. C., Friedman, R., Parker, J. D., Tofler, G. H., Jimenez, a H., Muller, J. E., … Stone, P. H. (1994). Use of skin conductance changes during mental stress testing as an index of autonomic arousal in cardiovascular research. *American Heart Journal*, *128*(6 Pt 1), 1170–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7985598

Joëls, M., Pu, Z., Wiegert, O., Oitzl, M. S., & Krugers, H. J. (2006). Learning under stress: how does it work? *Trends in Cognitive Sciences*, *10*(4), 152–8. doi:10.1016/j.tics.2006.02.002

Kappeler-Setz, C., Gravenhorst, F., Schumm, J., Arnrich, B., & Tröster, G. (2011). Towards long term monitoring of electrodermal activity in daily life. *Personal and Ubiquitous Computing*, *17*(2), 261–271. doi:10.1007/s00779-011-0463-4

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The "Trier Social Stress Test" - A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. *Neuropsychobiology*, *28*, 76–81.

Kivikangas, J. M., Chanel, G., Cowley, B., Ekman, I., Salminen, M., Jarvela, S., & Ravaja, N. (2011). A review of the use of psychophysiological methods in game research. *Journal of Gaming & Virtual Worlds*, *3*(3), 181–199. doi:10.1386/jgvw.3.3.181

Lenneman, J. K., & Backs, R. W. (2009). Cardiac autonomic control during simulated driving with a concurrent verbal working memory task. *Human Factors*, *51*(3), 404–418. doi:10.1177/0018720809337716

Loggia, M. L., Juneau, M., & Bushnell, M. C. (2011). Autonomic responses to heat pain: Heart rate, skin conductance, and their relation to verbal ratings and stimulus intensity. *Pain*, *152*(3), 592–598. doi:10.1016/j.pain.2010.11.032

London, H., Schubert, D. S., & Washburn, D. (1972). Increase of autonomic arousal by boredom. *Journal of Abnormal Psychology*, *80*(1), 29–36. doi:10.1037/h0033311

Mandryk, R. L., & Atkins, M. S. (2007). A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies*, *65*(4), 329–347. doi:10.1016/j.ijhcs.2006.11.011

Mandryk, R. L., Inkpen, K. M., & Calvert, T. W. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & Information Technology*, *25*(2), 141–158. doi:10.1080/01449290500331156

McLeod, D. R., Hoehn-Saric, R., & Stefan, R. L. (1986). Somatic symptoms of anxiety: comparison of self-report and physiological measures. *Biological Psychiatry*, *21*(3), 301–310. doi:10.1016/0006-3223(86)90051-X

Mehler, B., Reimer, B., Coughlin, J. F., & Dusek, J. A. (2010). Impact of Incremental Increases in Cognitive Workload on Physiological Arousal and Performance in Young Adult Drivers. *Transportation Research Record: Journal of the Transportation Research Board*, *2138*(-1), 6–12. doi:10.3141/2138-02

Miller, L. M., Chang, C. I., Wang, S., Beier, M. E., & Klisch, Y. (2011). Learning and motivational impacts of a multimedia science game. *Computers & Education*, *57*(1), 1425–1433. doi:10.1016/j.compedu.2011.01.016

Miller, S. M. (1979). Controllability and human stress: method, evidence and theory. *Behaviour Research and Therapy*, *17*(4), 287–304. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/486033

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist*, *38*(1), 63–71. doi:10.1207/S15326985EP3801_8

Ravaja, N., Saari, T., Salminen, M., Laarni, J., & Kallinen, K. (2006). Phasic Emotional Reactions to Video Game Events : A Psychophysiological Investigation, (June 2014), 37–41. doi:10.1207/s1532785xmep0804

Ravaja, N., Turpeinen, M., Saari, T., Puttonen, S., & Keltikangas-Järvinen, L. (2008). The psychophysiology of James Bond: phasic emotional responses to violent video game events. *Emotion (Washington, D.C.)*, *8*(1), 114–20. doi:10.1037/1528-3542.8.1.114

Salehi, B., Cordero, M. I., & Sandi, C. (2010). Learning under stress: the inverted-U-shape function revisited. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *17*(10), 522–30. doi:10.1101/lm.1914110

Setz, C., Arnrich, B., Schumm, J., Marca, R. La, Troster, G., & Ehlert, U. (2010). Discriminating Stress From Cognitive Load Using a Wearable EDA Device. *IEEE Transactions on Information Technology in Biomedicine*, *14*(2), 410–417.

Shi, Y., Ruiz, N., Taib, R., Choi, E. H. C., & Chen, F. (2007). Galvanic Skin Response (GSR) as an Index of Cognitive Load. In *CHI EA '07 CHI '07 Extended Abstracts on Human Factors in Computing Systems* (pp. 2651–2656).

Susi, T., Johannesson, M., & Backlund, P. (2007). *Serious Games – An Overview*.

T-Xchange. (n.d.). Retrieved from http://www.txchange.nl/about-t-xchange/

Tarvainen, M. P., Niskanen, J.-P., Lipponen, J. a, Ranta-Aho, P. O., & Karjalainen, P. a. (2014). Kubios HRV--heart rate variability analysis software. *Computer Methods and Programs in Biomedicine*, *113*(1), 210–20. doi:10.1016/j.cmpb.2013.07.024

Van Dooren, M., de Vries, J. J. G. G. J., & Janssen, J. H. (2012). Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology and Behavior*, *106*(2), 298–304. doi:10.1016/j.physbeh.2012.01.020

Zyda, M. (2005). From Visual Simulation to Virtual Reality to Games. *Computer Society*, *38*(9), 25–32.

APPENDIX A: INFORMED CONSENT FORM

Informatiebrochure

Beste lezer,

In deze brief willen we je informeren over het onderzoek waarvoor je je hebt aangemeld. Het experiment vindt plaats op ...-...-..., in ruimte B219C van de Cubicus. In het voorgestelde onderzoek, getiteld "Serious games en Electrodermal Activity" wordt de huidgeleiding, de hartslag en in hoeverre je gestrest bent gemeten en worden er dilemma's via een computerspel aangeboden. Doel van het onderzoek is vast te stellen of de dilemma's' die worden aangeboden in het spel van invloed zijn op de huidgeleiding en de hartslag. Met de resultaten van het onderzoek zou het spel hopelijk aangepast kunnen worden zodat de spelers 'optimaal aroused' zijn tijdens het spelen ervan. Het doel daarvan is om een zo realistisch mogelijke dilemma situatie te creëren.

Om de huidgeleiding en de hartslag te meten wordt gebruik gemaakt van sensoren die verwerkt zitten in een armband voor om de pols. Deze sensoren meten door een heel klein stroompje de geleiding van de huid. Dit stroompje is nog kleiner dan statische geleiding, en je voelt hier helemaal niets van. Verder wordt er tijdens het spel af en toe gevraagd hoe gestrest je je voelt. Na het spel wordt hier ook nog informatie over gevraagd in een vragenlijst.

In het onderzoek wordt je als proefpersoon geconfronteerd met een aantal dilemma's die studie gerelateerd zijn. Er zijn geen goede en foute antwoorden. Aan het eind van het spel krijg je feedback over de door jou gemaakte keuzen. In het spel zitten er een aantal personen aan tafel die jou informatie verschaffen over de situatie omtrent het dilemma. Door deze informatiebronnen te gebruiken kan je voor jou een goed beeld krijgen van de situatie en een beslissing nemen. Tijdens het spel wordt er gevraagd hoe gestrest je je voelt. Je kan dit aangeven op een schaal van 3 (groene smiley is niet gestrest, gele smiley is een beetje gestrest en de rode smiley is erg gestrest). Voor en na het spel zal je eerst ontspannende geluiden horen. Deze geluiden zijn er om voor een baseline te zorgen. Zo kan de stress die je misschien had voor het onderzoek even zakken. Aan het eind van het onderzoek zal er nog een vragenlijst gegeven worden met wat vragen over de stress tijdens het spel en vragen over demografische gegevens.

Verder is het relevant te weten dat de opstelling van alle apparatuur volstrekt veilig is, en dat de onderzoeker over ruime ervaring met deze opstelling en dergelijke stimuli beschikt.

Voor deelname aan het experiment is het van belang dat je geen medische problemen (pacemaker) hebt en dat je beschikt over een goed gezichtsvermogen en begrip van de Nederlandse taal. Verder kun je te allen tijde besluiten om tijdens het experiment te stoppen zonder dat dit voor jou consequenties heeft, waarbij ook geen reden aangegeven hoeft te worden. Tot dan toe verdiende vergoedingen worden dan ook gewoon (naar rato van de deelname) uitgekeerd. Verder kun je ook na afloop van het onderzoek, tot 24 uur daarna, alsnog besluiten *dat je gegevens niet verder mee worden genomen in het onderzoek. Ander relevante aspecten zijn dat met je gegevens op een vertrouwelijke wijze wordt omgegaan, en dat anonimiteit van je gegevens is gewaarborgd en dat deze nooit aan derden zonder jouw toestemming zullen worden verstrekt.*

Het experiment zal maximaal 1 uur duren en hiervoor ontvang je een beloning, 1 credit. Het is relevant te weten dat het merendeel van de proefpersonen deelname aan soortgelijke experimenten zeer interessant vindt.
Na afloop van het volledige onderzoek kun je, indien je dat wenst, middels een debriefing over de verkregen resultaten op de hoogte worden gesteld.

Met hartelijke groet,

Coördinator:
Dr. Matthijs Noordzij
Cubicus B333. Cognitieve Psychologie en Ergonomie Faculteit Gedragswetenschappen Universiteit Twente

Tel: 053-489 2589

E-mail: m.l.noordzij.utwente.nl

Proefleider/Onderzoeksasssitent:
Joyce Leussink
Calslaan 60-102, 7522 MG Enschede
Tel: 0657931968
E-mail: j.leussink-1@student.utwente.nl

Informed Consent

Ik verklaar hierbij op voor mij duidelijke wijze te zijn ingelicht over de aard en methode van het onderzoek, zoals uiteengezet in de bovenstaande informatiebrochure 'Serious games en Electrodermal Activity'. Mijn vragen zijn naar tevredenheid beantwoord. Ik stem geheel vrijwillig in met deelname aan dit onderzoek. Ik behoud daarbij het recht deze instemming weer in te trekken zonder dat ik daarvoor een reden hoef op te geven en besef dat ik op elk moment mag stoppen met het experiment. Indien mijn onderzoeksresultaten gebruikt zullen worden in wetenschappelijke publicaties, dan wel op een andere manier openbaar worden gemaakt, zal dit volledig geanonimiseerd gebeuren. Mijn persoonsgegevens zullen niet door derden worden ingezien zonder mijn uitdrukkelijke toestemming. Als ik nog verdere informatie over het onderzoek zou willen krijgen, nu of in de toekomst, kan ik me wenden tot Joyce Leussink.
Voor eventuele klachten over dit onderzoek kunt u zich wenden tot de secretaris van de Commissie Ethiek van de faculteit Gedragswetenschappen van de Universiteit Twente, mevr. J. Rademaker (telefoon: 053-4894591; e-mail:j.rademaker@utwente.nl, Postbus 217, 7500 AE Enschede).

Aldus in tweevoud getekend:

……………………………	*……………………………*
Naam proefpersoon	*Handtekening*

Ik heb toelichting verstrekt op het onderzoek. Ik verklaar mij bereid nog opkomende vragen over het onderzoek naar vermogen te beantwoorden.'

……………………………	*……………………………*
Naam onderzoeker	*Handtekening*

APPENDIX B: SCENARIO 'STUDY CHOICES'

Target group (players):

The target group for this scenario are students, because they often have problems with time management tasks like planning and making a schedule for their homework.

Training environment:

This scenario is made for research purposes, but can be used in the time management training that is given on universities.

Learning goals:

The learning goals from this scenario are giving the students more insight in their behaviour among the choices they have to make in their study. These choices can be between social life or doing homework.

Actors:

Role	Personality description	Picture
Study advisor	The study advisor gives the student advice about their study and makes them think about the consequences that choices have for their study.	
Mother	The mother wants the best for her child. She wants that the student is doing well in school but also does not get stressed.	
Friend	The friend of the student wants to have fun, and thinks that the student should have more fun instead of studying.	
Chairman association	The chairman of the association wants to get a new board, and thinks that the student is a good candidate for this job.	

Teacher	The teacher wants that the student gets a good grade for his course. He thinks that his course is the most exiting course within the study.	
Your conscience	The conscience of the students speaks out the inner thoughts of the student.	

Scenario

Title: Study choices.

Duration: 6 minutes (fast version) or 20 minutes (slow version).

The dilemmas from the fast version will be written out, the only difference with the slow version is that some figures give information twice and that the information pop-up time is shorter.

Scenario description (Dutch):

Je bent een student aan de Universiteit Twente. Op dit moment gaat het niet zo goed met je studie. Je stelt vaak opdrachten uit en vindt het lastig om je tijd goed in te plannen. Vorig jaar kwam je zelfs studiepunten te kort om door te mogen gaan met je studie. Maar na een goed gesprek met je studie adviseur is besloten dat je over mag gaan, indien je meer inzet toont en al je openstaande herkansingen in een keer haalt. Daarnaast heb je een bijbaan om je studie en huur te kunnen bekostigen en ben je lid van een sport/cultuur-vereniging. Bij deze vereniging ben je elke week te vinden om te trainen / oefenen. Je bent dus erg actief. Je bent zelfs gevraagd of je het komende studiejaar het bestuur in wilt.

Dilemma 1

Title: The library.

Situation description (Dutch): Een nieuw studiejaar is begonnen en je bent begonnen aan een nieuw vak. Een studiegenoot van je wil graag de UB in met je om samen te studeren.

Dilemma question: Will you join your classmate with studying in the library?

Role	**Advice**	**Information**
 Study advisor		**Title:** Goed begin.
	Advice: Yes	**Text:** Ik hoorde dat je studiegenoot erg gemotiveerd is, en die motivatie kan jij ook wel gebruiken.
	Time: 12	**Time:** 4
 Mother		**Title:** Slim.
	Advice: Yes	**Text:** Slim idee van je studiegenoot, en voor het bestuur heb je toch helemaal geen tijd?
	Time: 4	**Time:** 8
 Friend		**Title 1:** Overdreven. **Title 2:** Overdreven 2.
	Advice: No	**Text 1:** Ik hoorde dat het oude bestuur kennis met je wilt maken en uitleg wil geven. Dat lijkt me wel erg zinvol. **Text 2:** Activisme bij een vereniging staat wel goed op je CV!
	Time: 8	**Time 1:** 4 **Time 2**: 12
 Chairman association		**Title 1:** Hard nodig. **Title 2**: Beslissing.
	Advice: No	**Text 1:** We spreken deze week elke avond af, dus hopelijk heb je niets anders gepland! **Text 2:** Over een paar weken moet je beslissen of je in het bestuur wilt gaan.
	Time: 15	**Time 1:** 10 **Time 2**: 15
		**Title:** Eigen keus.
	Advice: Yes	**Text:** Voor dit vak staat veel tijd gepland dat jullie in zelfstudie moeten steken. Anders

Teacher		wordt het wel wat lastig om dit vak te behalen.
	Time: 16	**Time:** 0
		**Title:** Erg leuk.
Your Conscience	**Advice:** No	**Text:** Als ik een jaar bestuur doe ontwikkel ik mezelf natuurlijk wel... en het is erg leuk!
	Time: 0	**Time:** 12

Dilemma 2

Title: On a visit.

Situation description (Dutch): De tweede week van het kwartiel is aangebroken. Je moeder belt je op met de vraag of je zin hebt om een dag op visite te komen en mee te blijven eten.

Dilemma question: Are you going to visit your mother?

Role	**Advice**	**Information**
		**Title:** OV.
Study advisor	**Advice:** No	**Text:** Je vertelde dat de reis naar je moeder 2 uur duurt met het OV, voor een enkele reis. Het neemt dan wel heel veel tijd in beslag.
	Time: 2	**Time:** 12
		**Title 1:** Ziek. **Title 2:** Alstjeblieft.
Mother	**Advice:** Yes	**Text 1:** Ik ben al een tijdje erg ziek en heb je al zo lang niet gezien. Kom je alstjeblieft langs? **Text 2:** Alstjeblieft? Ik zou het echt heel fijn vinden als je langs kon komen.
	Time: 8	**Time 1:** 8 **Time 2:** 10
		**Title:** Ontspannen.
	Advice: Yes	**Text:** Een dag even ontspannen bijpraten bij je

Friend		moeder kan vast geen kwaad denk ik.
	Time: 5	**Time:** 8
		**Title:** Druk.
Chairman association	**Advice:** No	**Text:** Ik dacht dat je zoveel te doen had voor je studie? Krijg je dat dan wel af?
	Time: 10	**Time:** 3
		**Title:** Verslag.
Teacher	**Advice:** No	**Text:** Volgende week moeten jullie een verslag af hebben, heb je daar dan nog wel genoeg tijd voor?
	Time: 4	**Time:** 7
		**Title:** Familie.
Your Conscience	**Advice:** Yes	**Text:** Ik vraag me al een tijd af hoe het met haar gaat, ik heb haar al zo lang niet gezien.
	Time: 0	**Time:** 2

Dilemma 3

Title: The wedding.

Situation description (Dutch): Je weet het al een tijdje, zaterdag is de bruiloft van een goede vriend. De bruiloft zal de hele dag duren want je bent uigenodigd voor het middagprogramma en het diner met het afsluitende feest.

Dilemma question: Will you go to the wedding?

Role	**Advice**	**Information**
		**Title:** Te weinig tijd.
	Advice: No	**Text:** Als je naar de bruiloft gaat kan je de rest van de dag niet meer aan je studie besteden.

Study advisor	**Time:** 8	**Time:** 11
 Mother		**Title:** Wees verstandig.
	Advice: No	**Text:** Wees verstandig, je vriend heeft vast wel begrip voor de situatie.
	Time: 4	**Time:** 8
 Friend		**Title:** Belangrijk.
	Advice: Yes	**Text:** Kom op, mijn bruiloft staat al maanden gepland. Je kan het niet maken om niet te gaan.
	Time: 8	**Time:** 4
 Chairman association		**Title:** Afspraak.
	Advice: No	**Text:** Binnen de vereniging houden we er ook niet van als mensen hun afspraak niet nakomen.
	Time: 0	**Time:** 0
 Teacher		**Title:** Deadline.
	Advice: No	**Text:** Ik heb jullie vorige week gezegd dat er maandag een deadline is voor het verslag. Ik zou er als ik jou was wel genoeg tijd in steken.
	Time: 12	**Time:** 3
 Your Conscience		**Title 1:** Vriendschap. **Title 2:** Tijd tekort.
	Advice: Yes	**Text 1:** Zou hij me dit kunnen vergeven? **Text 2:** Ik weet bijna zeker dat als ik ga, ik tijd tekort kom..
	Time: 10	**Time 1:** 1 **Time 2:** 14

Dilemma 4

Title: The cleaning.

Situation description (Dutch): De tentamenweek komt in zicht en je hebt gepland om deze dag te studeren. Je zit aan je bureau maar kan je niet concentreren. Je kamer is ook zo'n rotzooi, die moet echt eerst schoongemaakt worden voordat je kan beginnen. En de afwas, het vuilnis, en nu je toch bezig bent kan je net zo goed je kast gaan opruimen.
Dilemma question: Will you clean your room today?

Role	**Advice**	**Information**
 Study advisor		**Title:** Planning.
	Advice: No	**Text:** Het is belangrijk dat je voor jezelf een planning maakt en je daar ook aan houdt. Het schoonmaken zou je kunnen inplannen, zorg dan wel dat je er niet langer aan besteedt dan gepland.
	Time: 6	**Time:** 9
 Mother		**Title:** Typische student.
	Advice: No	**Text:** Terwijl je moet leren gaan schoonmaken is nou weer typisch studentengedrag!
	Time: 5	**Time:** 4
 Friend		**Title:** First things first.
	Advice: Yes	**Text:** Die rommel leidt alleen maar af, dat heb ik ook altijd. Schoonmaken werkt echt!
	Time: 1	**Time:** 8
 Chairman association		**Title:** Belangrijk.
	Advice: Yes	**Text:** Kom je vanvaond ook nog even langs de vereniging?
	Time: 10	**Time:** 10
		**Title:** Tentamen.
	Advice: No	**Text:** Het is belangrijk dat je dit vak haalt, schoonmaken kan je ook na je tentamen doen.

Teacher	**Time:** 3	**Time:** 1
		**Title 1:** Wel zo fijn. **Title 2:** Wel zo fijn2.
Your Conscience	**Advice:** Yes	**Text 1:** Wanneer alles lekker fris ruikt kan ik me volledig op mijn studie concentreren! **Text 2:** Mijn handen beginnen al te jeuken, ik heb zin om aan de slag te gaan!
	Time: 11	**Time 1:** 2 **Time 2:** 11

Dilemma 5

Title: The board.

Situation description (Dutch): Je hebt nu een paar weken de tijd gehad om na te denken over de bestuursfunctie waarvoor je gevraagd bent. Het moment is gekomen waarop je een beslissing moet nemen of je je actief in wilt zetten voor deze vereniging.

Dilemma question: Will you join the board of your association?

Role	**Advice**	**Information**
		**Title:** Belangrijk3.
Study advisor	**Advice:** No	**Text:** Ga eens bij jezelf na wat nou het belangrijkste is dit jaar...
	Time: 7	**Time:** 6
		**Title:** Herkansing.
Mother	**Advice:** No	**Text:** Als het toch drukker wordt dan verwacht is er een kans dat je volgend kwartiel een extra tentamen hebt.
	Time: 6	**Time:** 9

 Friend		**Title:** Vertrouwen.
	Advice: Yes	**Text:** Ik hoorde dat de leden een goede voorzitter in je zien, ze hebben vertrouwen in je!
	Time: 9	**Time:** 0
 Chairman association		**Title:** Kleine vereniging.
	Advice: Yes	**Text:** Het is een kleine vereniging, dus zoveel tijd zal het niet in beslag nemen.
	Time: 2	**Time:** 10
 Teacher		**Title 1:** Studeren. **Title 2:** Studeren2
	Advice: Yes	**Text 1:** Ik heb gehoord dat een bestuursfunctie wel vaker gecombineerd wordt met de studie, maar het is en blijft je eigen keus. **Text 2:** Het is je eigen keus wat je belangrijker vind.
	Time: 5	**Time 1:** 1 **Time 2:** 12
 Your Conscience		**Title:** Slechte studie.
	Advice: No	**Text:** Het gaat op dit moment niet zo goed met mijn studie, dus ik moet hier wel veel tijd in steken..
	Time: 1	**Time:** 5

Dilemma 6

Title: The vacation

Situation description (Dutch): Je zou met je vriend na de tentamenweek op vakantie. Echter, je hebt vooruit gekeken voor het komende kwartiel en ontdekt dat de nieuwe studiestof al bekend is.

Dilemma question: Will you go on vacation with your friend?

Role	Advice	Information
 Study advisor		**Title:** Hoger cijfer.
	Advice: No	**Text:** Je vriend snapt het vast wel als jullie de vakantie opschuiven. En een hoger cijfer kunnen halen lijkt mij wel belangrijk.
	Time: 21	**Time:** 2
 Mother		**Title:** Belofte.
	Advice: Yes	**Text:** Ik heb je altijd geleerd dat je je wel aan je beloftes moet houden.
	Time: 9	**Time:** 8
 Friend		**Title:** Beloofd.
	Advice: Yes	**Text:** Je had beloofd dat we zouden gaan...
	Time: 15	**Time:** 5
 Chairman association		**Title:** Besluitvaardig.
	Advice: No	**Text:** Als je vooruit gaat werken heb je in dit kwartiel ook meer tijd over voor andere zaken naast je studie..
	Time: 2	**Time:** 10
 Teacher		**Title:** Voorbereiding.
	Advice: No	**Text:** Als je vooruit gaat werken ben je perfect voorbereid, daar pluk je rond de tentamenweek de vruchten van!
	Time: 5	**Time:** 3
		**Title:** Verdiend.
	Advice: Yes	**Text:** Vakantie, dat heb ik ook wel verdiend na

Your Conscience		deze lastige tentamenweek.
	Time: 11	**Time:** 6

APPENDIX C: QUESTIONNAIRE ABOUT SUBJECTIVE STRESS

Vragenlijst Serious games & Electrodermal Activity

Bedankt voor het deelnemen aan mijn onderzoek naar de invloed van een serious dilemma game op de huidgeleiding en hartslag van spelers.
Om de resultaten van het onderzoek goed te kunnen interpreteren heb ik nog wat demografische gegevens van je nodig en informatie over stress.

Demografische gegevens

1. Geslacht *NL/ DE*

- o Man
- o Vrouw

2. Leeftijd

- o 16 t/m 20
- o 21 t/m 25 .... jaar
- o 26 t/m 30
- o Ouder dan 30

3. In welk jaar van je studie zit je?

- o 1e jaar bachelor
- o 2e jaar bachelor
- o 3e jaar bachelor
- o 4e jaar bachelor (uitloop)
- o 1e jaar master
- o 2e jaar master
- o Anders namelijk: ....

Vragen over het spel en stress

Geef je score aan op een schaal van 1 tot 10.

4. In hoeverre was je gestresst voorafgaand aan het ondezoek?

Niet gestresst 1 2 3 4 5 6 7 8 9 10 heel erg gestresst

5. In hoeverre was je gestresst tijdens het spel?

Niet gestresst 1 2 3 4 5 6 7 8 9 10 heel erg gestresst

6. In hoeverre was je gestresst na het spelen van het spel?
Niet gestresst 1 2 3 4 5 6 7 8 9 10 heel erg gestresst

7. Wat vond je stressvol aan het spel?
- o De tijdsdruk.
- o De dilemma's waren lastig.
- o De dilemma's kwamen snel achter elkaar.
- o De informatie-items van de personen kwamen snel.
- o Anders namelijk:...

Algemene vragen over stress

8. Ben je snel gestresst?
- o Ja
- o Nee

9. Waardoor ben je snel gestresst?
- o Studie druk
- o Tentamens
- o Sociale verplichtingen
- o Teveel keuzes tegelijk moeten maken
- o Hobby's
- o Anders namelijk:...

Algemene vraag over het spel

10. In hoeverre kon je je inleven in onderstaande situatiebeschrijving?

'Je bent een student aan de Universiteit Twente. Op dit moment gaat het niet zo goed met je studie. Je stelt vaak opdrachten uit en vindt het lastig om je tijd goed in te plannen. Vorig jaar kwam je zelfs studiepunten te kort om door te mogen gaan met je studie. Maar na een goed gesprek met je studie adviseur is besloten dat je over mag gaan, indien je meer inzet toont en al je openstaande herkansingen in een keer haalt. Daarnaast heb je een bijbaan om je studie en huur te kunnen bekostigen en ben je lid van een sport/cultuur-vereniging. Bij deze vereniging ben je elke week te vinden om te trainen / oefenen. Je bent dus erg actief. Je bent zelfs gevraagd of je het komende studiejaar het bestuur in wilt.'

Helemaal niet 1 2 3 4 5 6 7 8 9 10 helemaal wel

APPENDIX D: MUSIC VIDEO BASELINE

https://youtu.be/aQxMO6Mt3Jc

This link leads to the music video that was used for the two baselines. The music was downloaded from YouTube and three different images of flowers were added so that the participants had to look at something that did not give too much arousal.

APPENDIX E: EXAMPLES BAD MEASUREMENTS HRV

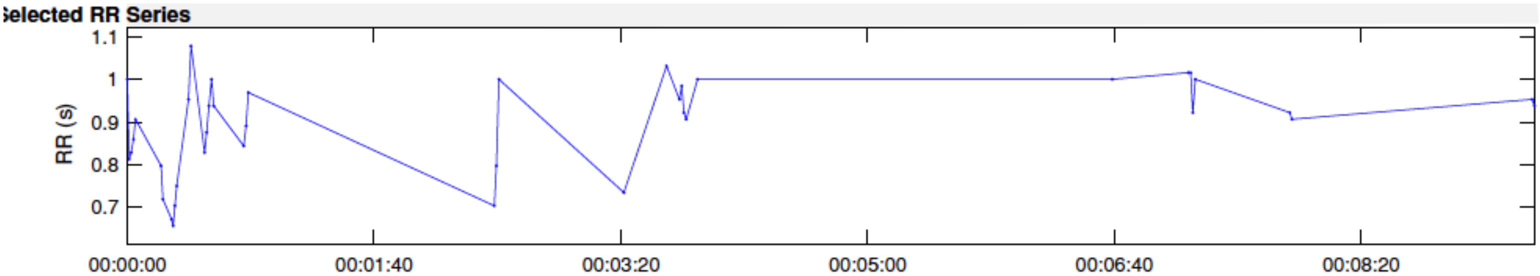

Figure 4. In this figure is shown an example of a dataset with very few measurement points. It shows that the Empatica E3 did not measured HR well for periods more than a minute.

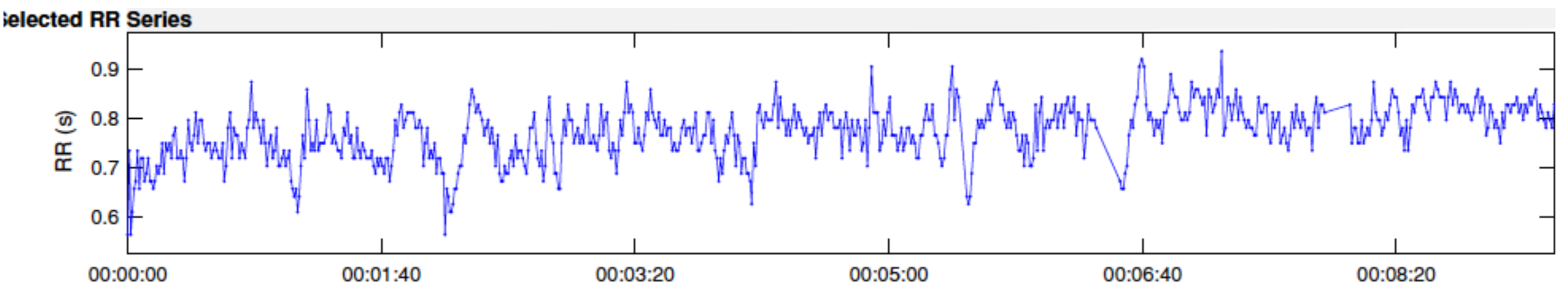

Figure 5. This figure shows an example of a correct dataset with enough measurement point.

APPENDIX F: DISTRIBUTION OF THE DEPENDENT VARIABLES

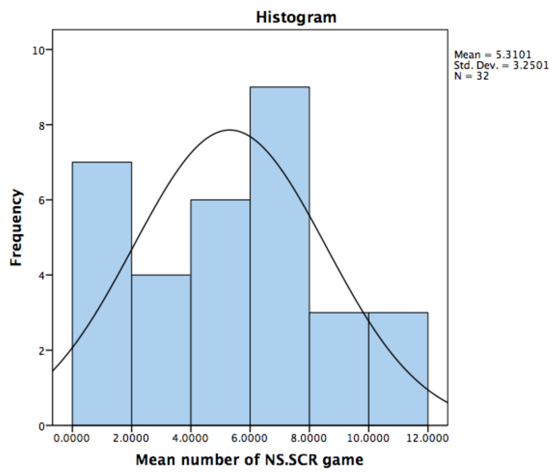

Figure 6. In this figure the distribution of the dependent variable ‘number of NS.SCR during the game’ is shown. This data is approximately normal distributed, so a Repeated measures Two-Way ANOVA could be performed.

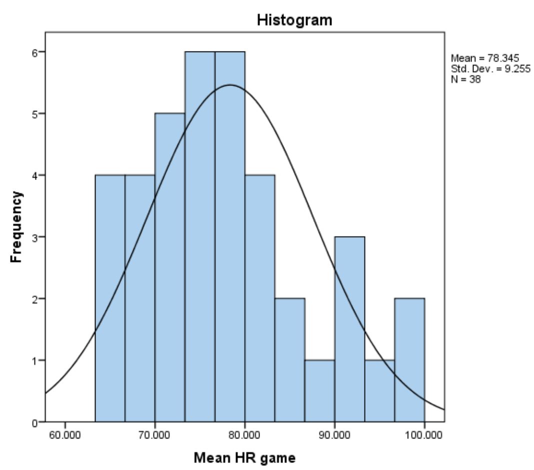

Figure 7. In this figure the distribution of the dependent variable ‘mean HR during the game’ is shown. This data is approximately normal distributed, so an Repeated measures Two-Way ANOVA could be performed.

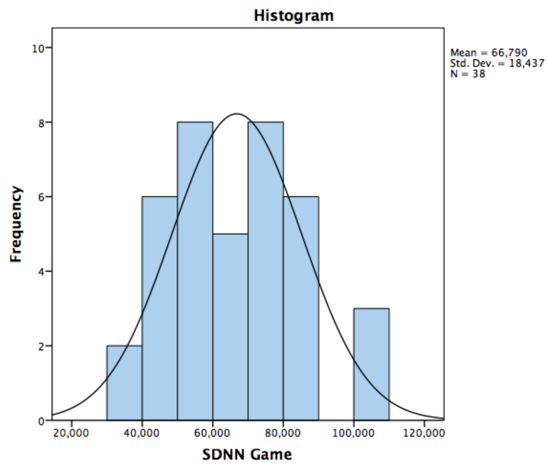

Figure 8. In this figure the distribution of the dependent variable 'mean SDNN during the game' is shown. This data is approximately normal distributed, so a Repeated measures Two-Way ANOVA could be performed.

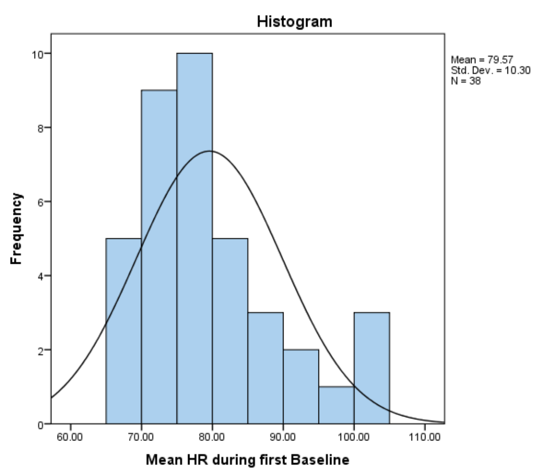

Figure 9. In this figure the distribution of the dependent variable 'Mean HR during the first baseline' is shown. This data is approximately normal distributed, so an independent samples t-test can be performed.

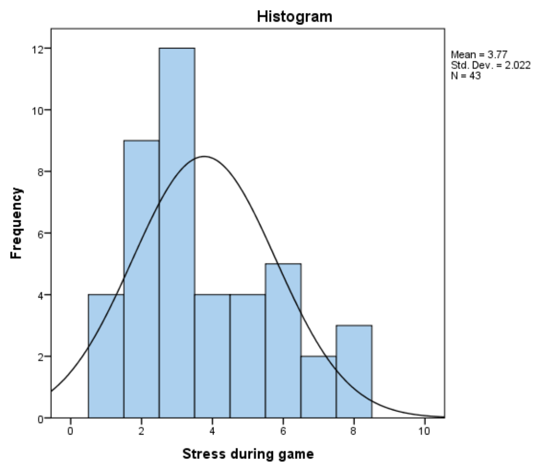

Figure 10. In this figure the distribution of the dependent variable 'experienced stress during the game' is shown. This data is skewed to the left, so an independent samples t-test could not be performed. A Mann-Whitney U test was performed instead.

APPENDIX G: GRAPHS AND TABLES FOR THE CORRELATION BETWEEN EDA, HR, AND HRV

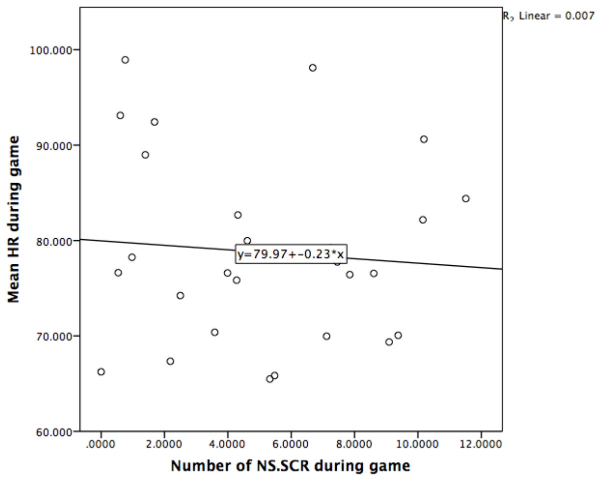

Figure 12a

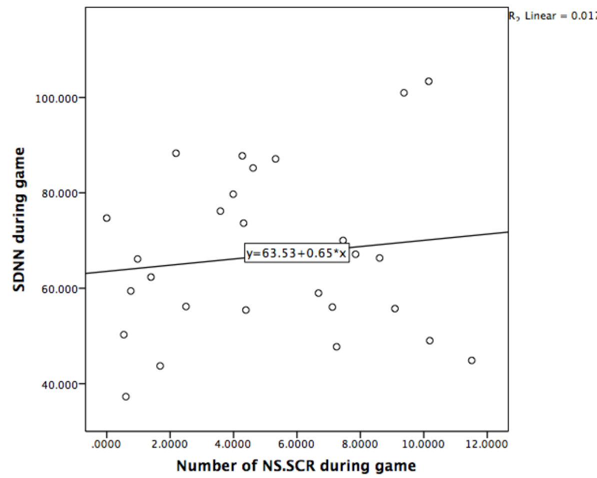

Figure 12b

Figure 12a and 12b: In figure 10a the correlation between EDA and HR is shown. It can be seen that when the number of NS.SCR increases, the HR decreases. In figure 10b it is visible that the variance increases when the number of NS.SCR also increases.

APPENDIX H: GRAPHS AND TABLES FOR THE CORRELATION BETWEEN THE PHYSIOLOGICAL AND EXPERIENCED STRESS

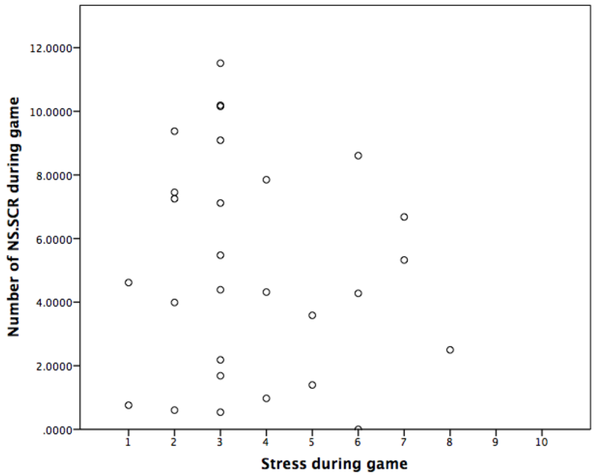

Figure 13a

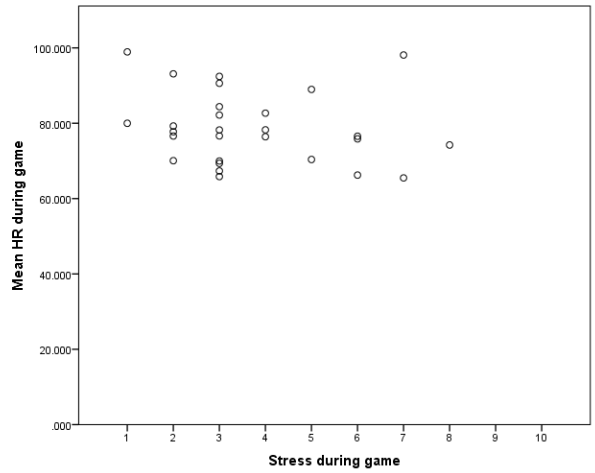

Figure 13b

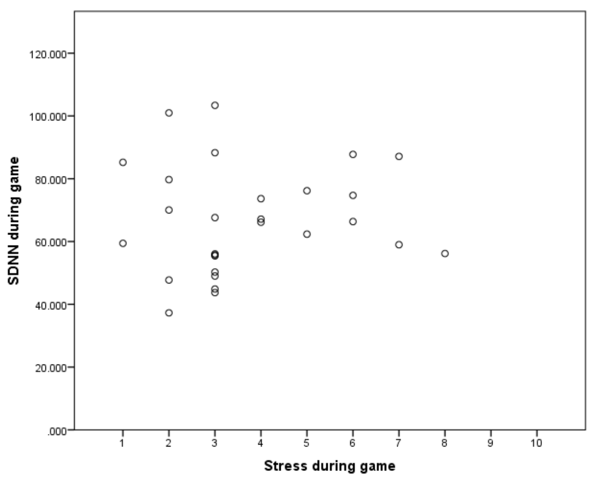

Figure 13c

Figure 13a, b, and c: In these figures the correlation between the subjective stress and the EDA, HR, and HRV is shown.

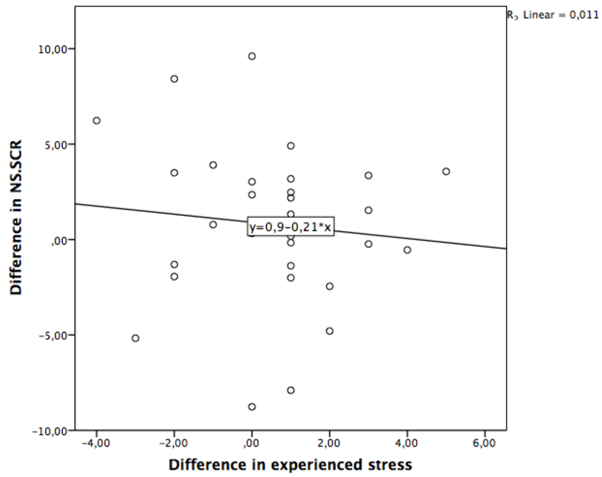

Figure 14a

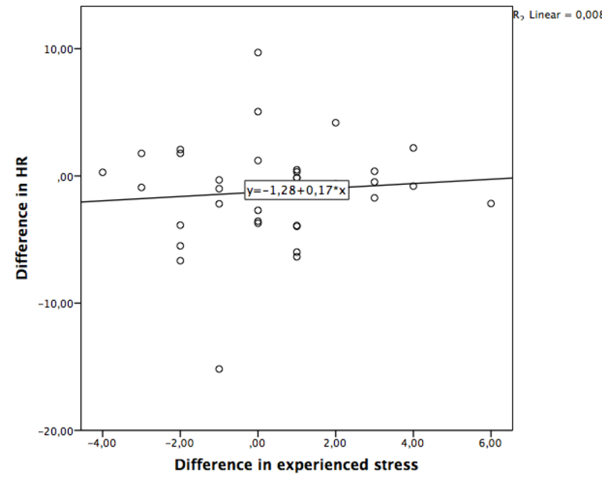

Figure 14b

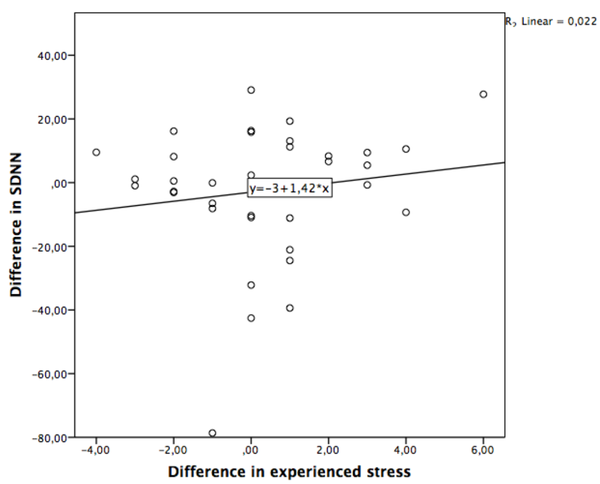

Figure 14c

Figure 14a, b, and c: In these figures the correlation between the difference in subjective stress and difference in the physiological measures is shown. No correlations were found.