Eindhoven University of Technology

MASTER

Doubling Doppel

the effect of a tactile heartbeat stimulus on arousal levels : a replication with critical review

van der Velden, G.P.W.

Award date:
2020

Doubling Doppel - The effect of a tactile heartbeat stimulus on arousal levels

A replication with critical review

Geert van der Velden

May 2020

Written as partial fulfillment of the requirements for the final degree of

Master of Science

in

Human-Technology Interaction

Supervisors:	Prof. Dr. J.H.D.M. Westerink
	Dr. D. Lakens
Assessor:	Dr. Ir. S. Schoenmakers

TABLE OF CONTENTS

TABLE OF FIGURES

ABSTRACT

Anxious moments seem to be ubiquitous in our busy society, with stress-related pathologies becoming more and more common. To counteract these stressors, wearables that promise to lower stress in a non-obtrusive way have entered the market.

This study focused on one such device: the Doppel. Doppel is a wrist-mounted device emulating heartbeat by ways of external, subtle vibration. We hypothesized that stress-induced arousal could be mitigated through applying this device by setting its external vibration to a frequency 20% lower than a person's resting heartrate. The proposed pathway to which this lowering of arousal could be achieved was through cardiac entrainment to the externally applied stimulus: lowering the heartrate and subsequently lowering perceived arousal. Another pathway is the degree of interoception that our participants exhibit, possibly influencing the amount of arousal they experienced.

We conducted a laboratory experiment where participants were fitted with the device vibrating at a frequency 20% lower than their resting heartrate while their heartrate, heartrate variability and skin conductance were recorded. Participants were faced with a stressful task which successfully induced arousal, after which they reported their subjective findings. Lastly, an interoceptive test was performed to determine their interoceptive ability.

Our results show no significant effect on the physiological and subjective measures of arousal. No heartrate entrainment was found and even the individuals scoring highest on interoception showed no significant beneficial effect. The conclusion of this study is that we found no support for our hypothesis that a externally applied haptic device induced an arousal-lowering effect, nor a cardiac entrainment effect.

Keywords: arousal, wearables, stress, entrainment, interoception

1 INTRODUCTION

1.1 GENERAL INTRODUCTION

In recent years, computational power, power consumption and reduction of electronics size has soared; enabling vast amounts of technology to be incorporated into our everyday lives. With sensors and actuators becoming more and more ubiquitous, smaller and cheaper, it is easier to gather data as well as actively steer behavior. One such development is the advent of *wearables*: sensors and actuators small enough so that they can be carried, varying from the size of a coin to full body gear. These *wearables* are enablers of several functions, from a more personal level like the Quantified Self (Swan, 2012), to bigger societal applications like assisted living (Rashidi & Mihailidis, 2013).

One practical application of wearable technology is the ability to influence and assist people in daily living, either consciously or subconsciously. Wearing sensors and actuators on your body can make tasks easier, give better insight into your health and influence your mental and somatic health, among others (Park & Jayaraman, 2003). A particular subset of the last category is wearable, calming devices like the Doppel, Touchpoint wristband and the Muse Headband. These relatively small apparatuses are utilized by consumers and patients to actively support them in achieving a more relaxed lifestyle, either by influencing people consciously or subconsciously over different modalities, see figure 1.

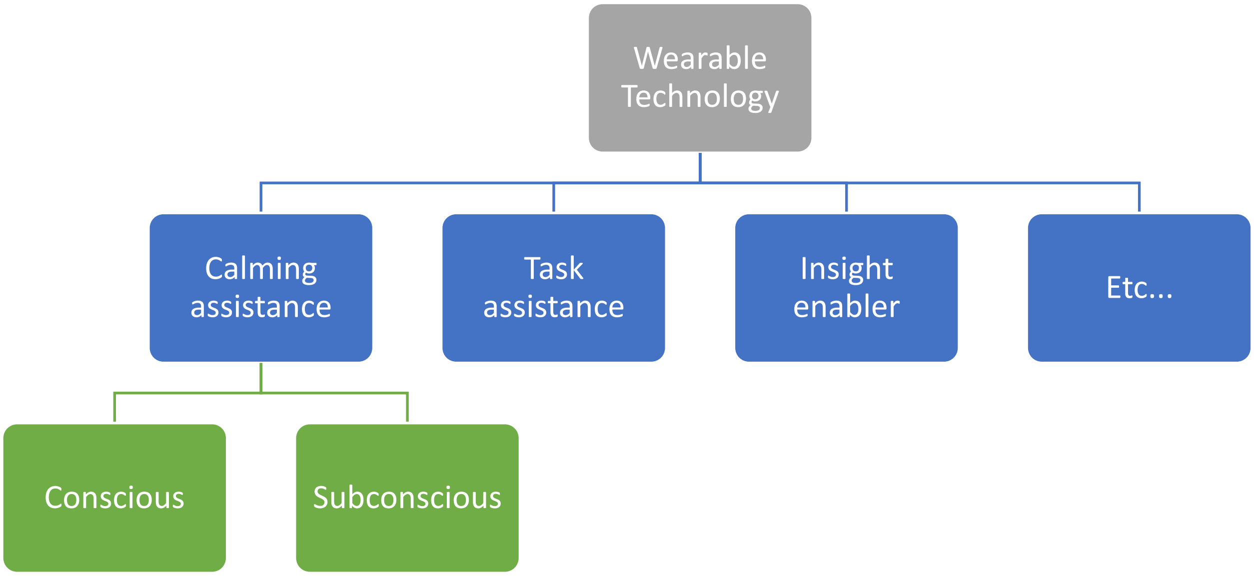


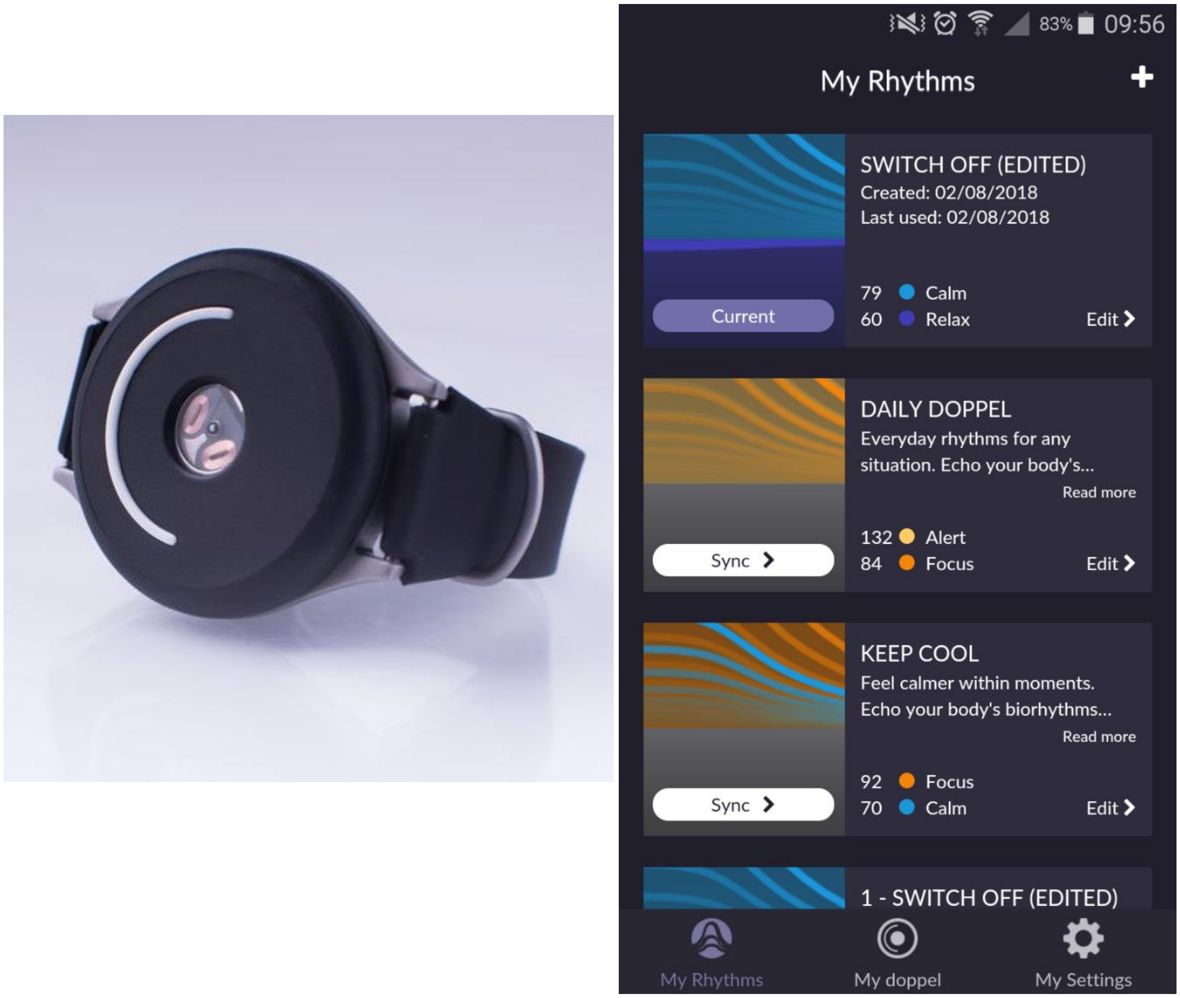
Figure 1: Brief overview of possible wearable subsets

The conscious way of influencing people is by means of notifications when certain actions need to be performed. An example is for support of medication adherence, hereby removing the stress associated with reminding oneself or forgetting medication altogether. These wearable devices prompt a user response that needs to be done actively by the user of the wearable (Serdaroglu, Uslu, & Baydere, 2015).

Another category is the one that we are most interested in in this thesis: the subconscious influencing wearable. Subconscious influencing can be done, as stated, over different modalities. Since wearables can be worn all over the body, different sensations in different places can be evoked through tactile, auditory or visual and other stimuli (Dijk & Weffers, 2013). These stimuli are used to evoke or inhibit certain behaviors or thought patterns, thereby improving the quality of life of consumers using the device (Gallace & Spence, 2010). A subset of the previously mentioned stimuli that is of interest for this thesis is tactile stimulation of people by using a heartbeat emulating device,possibly eliciting a lower amount of arousal in people when faced with a stressful task. The wrist-mounted Doppel is one such device that claims to do just that.

1.2 THE DOPPEL DEVICE AND THE DOPPEL STUDY

An example of a heartbeat emulating device and the focal point of this study is the *Doppel* vibrating watch, a device that emulates heartbeats by vibrating in customizable frequencies. The Doppel is a kickstarted device, capable of pairing wireless via Bluetooth to an app on one's phone. The app is able to set the frequency in which the device vibrates, the intensity in which it vibrates and toggle pre-determined rhythms. For use of the Doppel as a stress management device, it is advised to select a stimulation pattern that is below one's resting heartrate,



My Rhythms

SWITCH OFF (EDITED)
Created: 02/08/2018
Last used: 02/08/2018
79 Calm
60 Relax
Current
Edit

DAILY DOPPEL
Everyday rhythms for any situation. Echo your body's...
Read more
132 Alert
84 Focus
Sync
Edit

KEEP COOL
Feel calmer within moments. Echo your body's biorhythms...
Read more
92 Focus
70 Calm
Sync
Edit

1 - SWITCH OFF (EDITED)

Figure 2: The Doppel vibrating watch and the corresponding User Interface (UI)

Research concerning Doppel theorized that it influences the user by entrainment of the heart frequency (Phillips-Silver, Aktipis, & A. Bryant, 2010) with the frequency of Doppel's vibration, hereby heightening focus and lowering arousal(Azevedo et al., 2017). While the calming effect that Doppel has on users has been found while using its haptic vibration function, the precise reason why is still unknown since entrainment was not found in either studies conducted by Tsakiris (2017) and Azevedo et al. (2017).

For this thesis, we are particularly interested in the effects found by Azevedo et al. (Azevedo et al., 2017), where the central research questions of the original study were:

- *Can a person cope better with anxiety-induced arousal by usage of the Doppel device as opposed to people not using the device?*
- *Can the decrease in arousal be explained by ways of cardiac entrainment?*

In their study, they found that people experience a reduced amount of arousal when faced with a stressful task if a Doppel device was used set to a frequency 20% lower that their resting heart rate. They theorized that this would be due to entrainment of the heart frequency to the Doppel's frequency, however this effect was found to be insignificant. Nevertheless, an effect size of Hedges' $g = 0.54$ for the post-analysis independent t-test was found, indicating a rather large effect size that looks promising when studied further.

In this thesis we will look at the original study, attempt to replicate the results, extend it with extra data and critically review the original and replication study.

2 THEORETICAL BACKGROUND

In this section we will provide background to a number of topics that are relevant for our replication study. First we will elaborate on replication research, and describe how sequential analysis allows us to come to conclusions using a minimum number of participants. Then we will dive into the topics of stress, stress measurement, and stress reduction by HR entrainment, since the latter is what the Doppel device was designed to do. We will finish with a consideration of human interoception of (entrained) bodily parameters, and an elaboration of the research hypotheses of the present study.

2.1 REPLICATION RESEARCH

2.1.1 General

Replication is one of the hallmarks of science and has been a pillar of the empirical sciences for many years. It is based on the preposition that nature behaves lawfully and thus adheres to a certain set of rules, yielding the same results when repeated under the same set of circumstances (Dilworth, 1994). Based on this preposition, we could theoretically reproduce results from a well thought out and subsequently conducted experiment indefinitely (Hempel, 1968).

This theoretical standpoint generally does not fully hold up in practice. Depending on the type of research, a certain amount of variation is always to be expected, especially in the social sciences, where it is a difficult endeavor to keep circumstances constant (Freese & Peterson, 2017). In addition to these difficulties, social research is insufficiently accustomed to proper documentation of experimental practices and data; along with a slight aversion against publicizing (direct) replication research (Schmidt, 2009). These attitudes and practices have seen a significant improvement due to increased attention and awareness of the need of replication in the social sciences, culminating the Replication Movement.

2.1.2 The Replication Movement

The last couple of years, different subfields of science have seen a renewed interest and increase in reproducing past findings after discovering that some major studies do not hold up to scrutiny (Aarts et al., 2015). This movement has reached the psychological and other sciences, culminating in a combined effort to reproduce 100 prominent studies, all with significant results. The results were mixed, with replication effects being half of the original effects and 40% of the replicated studies showing a significant effect. The study has shed light on some errors in contemporary research and how to circumvent these, like a better documentation of used methods, higher powered studies, open data, heightened scrutiny on publication bias and a better insight into used analysis techniques.

2.1.3 Types of replication and their goals

Several types of replication are to be discerned in science, depending on the goal of the replication. Two definitions are frequently utilized (Schmidt, 2009):

- *Direct replication:* The replication of an experimental procedure;
- *Conceptual replication:* Repetition of a test of a hypothesis or a result of earlier research work with different methods.

With *direct replication*, a researcher is aiming at replicating the study as true as possible, hereby replicating hypothesized effects using a pre-described method and analysis. In *conceptual replication*, the effect or hypothesis is tested using a different method, thus deviating from the original research's work. Schmidt (2009) defined a number of points where *direct* or *conceptual* replication of a study might fit:

1. To control for sampling error (chance result);
2. To control for artifacts (lack of internal validity);
3. To control for fraud;
4. To generalize results to a larger or to a different population;
5. To verify the underlying hypothesis of the earlier.

Fraud and artifacts are reasons for a direct replication since replicating previous studies accurately allows for ways to eliminate the lack of internal validity or outright fraud. The other three reasons benefit from a more conceptual replication (Schmidt, 2009). In this type, a researcher is able to modify some aspects of the study like a larger or more diverse sample, or look at the underlying hypothesis and approach this slightly differently by modifying some variables.

A thoroughly executed replication adheres to a number of ground rules, ensuring that the results are trustworthy (Brandt et al., 2014).

1. Carefully defining the effects and methods that the researcher intends to replicate;
2. Following as precise as possible the methods of the original study (including participant recruitment, instructions, stimuli, measures, procedures, and analyses);
3. Having high statistical power;
4. Making complete details about the replication available, so that interested experts can fully evaluate the replication attempt (or attempt another replication themselves);

5. Evaluating replication results and comparing them critically to the results of the original study.

The above ingredients for a well-executed replication are utilized to generate robust and trustworthy results, thus making sure that scarce resources are used effectively when building upon past research findings (Burman, Reed, & Alm, 2010). Another way of ensuring that resources are efficiently allocated is determining when to stop current research. In this, sequential analysis of research design is essential.

2.2 SEQUENTIAL ANALYSIS IN RESEARCH

Research in itself is a costly endeavor, both in time and resources. In design a study, gather data and analyze these data, a significant amount of resources are necessary to finally draw reliable conclusions from a particular study. One way to be more efficient and lower the amount of resources is to incorporate *sequential analysis* into the research design, allowing researchers to lower the amount of data points and get an intermediate look at where the research is heading at preset stages of the research (Sanderson & Fisher, 1994).

Sequential analysis in studies has its basis in the ambition to perform high-powered studies, all the while controlling for an inflated *type-I error* (Streiner, 1993). When a researcher analyzes data repeatedly while performing data gathering, the risk of aforementioned error increases, hereby possibly complicating analysis or invalidating the findings altogether. When the results of a study are checked in the meantime, the inevitable lower sample size poses a challenge in keeping the study sufficiently powered, while repeated analysis increases the odds of finding a positive effect (Murayama, Pekrun, & Fiedler, 2014). To control for this issue, a number of methods can be employed.

Several methods exist to design a study using sequential analysis, depending on the expected intermediate results of the study (Lakens, 2014). If a study is designed for an theorized significant effect, an *a priori* power analysis needs to be conducted (Pocock, 1977). In this design, an effect size estimation is required. The analysis results in an *unconditional power*, namely the power the study would have upon completion with the full sample set. The early interim analysis creates the possibility to check the *conditional power* of the study and determine whether an extrapolation of the findings would yield satisfactory results, possibly terminating the data collection early. A sequential analysis *a priori* can also modify the predefined alpha-level by lowering it, therefor creating the possibility for researchers to check their data while data collection is not yet completed (McPherson & Armitage, 1971).

A study can also be examined in the interim for the estimated effect size. When the findings show that the effect size is well below the smallest effect size of interest (SESOI) with no indication that this effect size will grow with more data collection, data collection may be terminated (Harms & Lakens, 2018). Data collection may also be terminated if the effect size exceeds the original estimate, while still having a power level that justifies the data collection thus far. Alternatively, if the effect size seems to attain the desired effect size upon completion of the data collection, the study may be continued.

2.3 STRESS

2.3.1 Stress: Physiology

The human body contains an elaborate nervous system that elicits control over every system in the body. Main components of the human nervous system are the central and peripheral system, with the central nervous system consisting of the brain and the spinal cord, and the peripheral nervous system consisting of the nerves and ganglia outside of the central nervous system (Mai & Paxinos, 2012). Both systems are responsible for maintaining *homeostasis*, i.e. the process of balancing the complex mechanisms inside our bodies by which a physiological equilibrium is achieved.

The peripheral nervous system can be divided further into the Somatic Nervous System (SNS), responsible for somatosensory and sensory systems, and the Autonomous Nervous System (ANS), regulating various physiological aspects of our bodies. The ANS is of particular interest in this research since the two systems that it comprises, namely the parasympathetic and sympathetic systems, are responsible for *rest and digest* and *fight or flight* responses, respectively (Mai & Paxinos, 2012).

When faced with a *stressor*, an influence that may upset homeostasis, our bodies elicit a *stress response* to preserve homeostasis during a time of immediate stress (Chrousos & Gold, 1992). Here the central and peripheral nervous system respond in a number of ways to restore homeostasis by freeing up resources to deal with this imbalance. The Central Nervous System (CNS) has the following functions in this response:

- Facilitate arousal, attention, cognition and aggression;
- Inhibits reproductive, growth and feeding functions;
- Activates counter-balance regulation feedback loops.

The Peripheral Nervous System (PNS) reacts with the following actions, among others:

- Increases oxygenation of organs and muscular tissue;

- Provides extra nutrition to brain, heart and skeletal muscles;
- Increases metabolism;
- Facilitate arousal;
- Immunosuppression. (Chrousos, 2009)

All of these responses temporarily disturb homeostasis in order to combat the stressor, with the intention of ultimately trying to revert back to homeostatic values when the stressor has subsided and the stress response in both the CNS and PNS has subdued.

The entire stress response is a complex interplay of cognitive appraisals, predispositions towards previous stressors, neurological triggering mechanisms and neural, endocrine and neuroendocrine responses. See figure 3 for a slightly redacted modeling of the human stress response (Everly, & Lating, 2012).

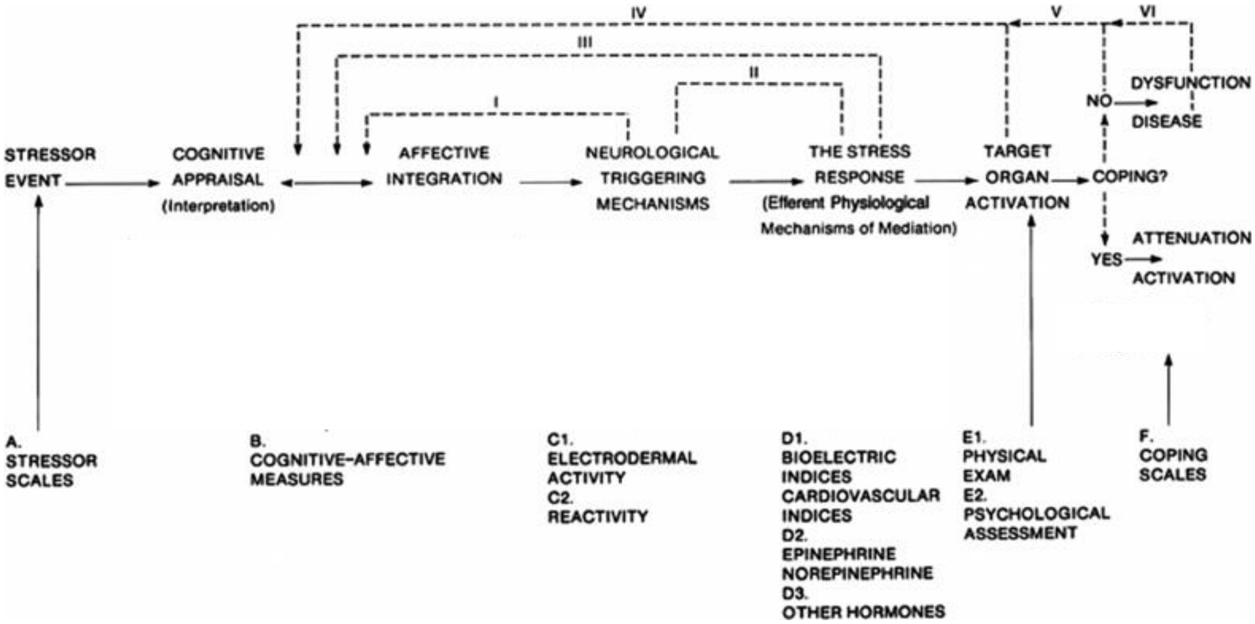


Figure 3: The Human Stress Response (Everly, & Lating, 2012)

Figure 3 shows the pathway of the body's stress-response, starting with the stressor event (A) and continuing on to the cognitive-affective measures (B). Here, the appraisal of the stressor takes place, taking into account several personal aspects like previous experiences and biological differences, among others. The stressor is interpreted and affectively integrated, leading to a neurological response, shown in C1 and C2. The stress response results in the secretion of several hormones (D), which are picked up by the relevant targeted organs (E). Section F finally shows the resulting outcome, leading to either coping with the stressor, or non-coping, possibly leading to adverse results. In between these stages, various feedback loops exist, making the process not entirely linear and dependent on state of the body, state of mind and other parameters.

All of these actions combined allow the body to cope with the stressor for a short period of time. When the stress response is prolonged, homeostasis is threatened and allostasis takes its place. The *allostatic load*, a term for an accumulation of allostatic effects over time, when prolonged can have negative effects on mind and body, like obesity, sleep-disorders, endocrinological disorders and depression (Joëls & Baram, 2009).

Depending on the type and intensity of stress, it is possible to measure an alteration of certain body parameters like decreased heart rate variability, increased cortisol levels and increased sweat production, among others (McGuigan & Andreassi, 1981). Those temporary bodily responses are used in measuring the effect that a certain stressor has in an individual, a feature that will be utilized later in this paper.

2.3.2 Stress: Psychology and social stress

Psychological factors of stress are at the root of the stress reaction, starting at the cognitive appraisal of a stressor, going through the body response and ending with possible coping, eventually resulting in possible allostatic load (Everly, & Lating, 2012), see figure 3. The way we appraise a stressor is the foundation upon which the rest of the stress response is built; this appraisal is shaped by previous experiences, genetics and overall mood, among others (Smith, 1989). There is a significant variety in what is considered a stressful event across cultures and individuals, with some specific events eliciting stress regardless of individual differences. One example is a clear and present danger causing the fight or flight response (McCarty, 2016). What is of particular interest in this paper however is *social stress* not necessarily induced by clear and present danger. The main reason why we are interested in this phenomenon is due to social stress being ubiquitous and the most important source of stress in modern day society, heightening the importance of dealing with this specific stressor.

Social stress occurs whenever people are subjected to a task or situation to which they exhibit a stress reaction due to fear of evaluation. Their emotional homeostasis is temporarily thrown off balance while the body is determining how to allocate resources to handle the stressor. When the previous action is finished, the emotional reappraisal takes place to determine that homeostatic balance has been restored (Hellhammer & Schubert, 2012). When persons are subjected to such a stressor, many of the same physiological responses are elicited as when faced with a stress reaction of a different nature, namely:

- Increased heart rate (measured via ECG);

- Decreased heart rate variability (measured via ECG);
- Increased cortisol levels (measured through blood or saliva tests);
- Increased sweat production (measured through Skin Conductance, (Pereira, Almeida, Cunha, & Aguiar, 2017).

Social stress can occur when people are faced with an instance of *social evaluation threat*, which occurs whenever people are faced with a situation that challenges their social standing or self-image. In human culture, social self-preservation is of utmost importance for survival in society. Whenever they are faced with a stressor that might endanger their social status, a stress response is elicited (Dickerson & Kemeny, 2004). A relatively simple way of inducing social stress is making people perform a public speaking task, either prepared well in advance or on the spot. A well-known paradigm in elicitation of social stress in the laboratory is the Trier social stress test (Kudielka, Wüst, Kirschbaum, & Hellhammer, 2007), a test where people are faced with a public speaking task. Several studies have shown that performing a speech or presentation in public creates a condition of *self-evaluative threat*, linked to social self-preservation (Westenberg et al., 2009). Recording the above listed parameters provides physiological data on the amount of stress that people experience, mainly recorded through Skin Conductance measurements and ECG's. These data can be appended by self-reported measurements like the State Trait Anxiety Inventory, model Y1 (STAI-Y1) or Brief Fear of Negative Evaluation (bFNE) scales (Carleton, McCreary, Norton, & Asmundson, 2006; Julian, 2011). Both of these scales measure the self-reported degree to which people are feeling stressed (STAI-Y1) or what their attitude towards public speaking is (bFNE).

2.4 PHYSIOLOGICAL MARKERS IN REACTION TO STRESS

In response to a stressor the human body elicits several reactions to combat the stressor, as we have described earlier. For this paper, we are interested in two major organs that alter their behavior in a phase of stress: the heart and the skin.

2.4.1 Heart rate and heart rate variability

The human heart is a muscle responsible for circulating blood through our vascular system. Control over the heart largely rests in the autonomous nervous system, where the parasympathetic and sympathetic systems regulate cardiac activity based on physical activity (Thayer, Åhs, Fredrikson, Sollers, & Wager, 2012). The dominant inhibitory influence of the parasympathetic system causes heart rate to favor energy conservation, resulting in high inter-beat (RR) intervals, while a higher sympathetic system response causes an increase in heart rate. Research by Saul (1990) showed that sympathetic effects on the heart are relatively slow, needing seconds to adapt to a new situation, whereas parasympathetic responses are much faster, requiring milliseconds to evoke a reaction (Saul, 1990). It is due to these differences in reflex times that the parasympathetic system is the primary reason for increased or decreased inter-beat variability. The parasympathetic system's fast reaction thus allows to finely control heart rate variability (HRV), to best accommodate various processes occurring in the body.

When faced with a prolonged or instantaneous stress response, the ability for the body to regulate dynamic autonomous processes is impaired, heartrate variability being a proxy for this impairment. (Prinsloo, Derman, Lambert, & Laurie Rauch, 2013). Having a high heart rate variability demonstrates the body's ability to anticipate goal-oriented behavior and control. Contrarily, people who report a prolonged or instantaneous period of stress have shown to have a lower resting heart rate variability as opposed to low self-reported stress people (Melzig, Weike, Hamm, & Thayer, 2009). In meta analyses, HRV has shown to be a prominent biomarker for prolonged or instantaneous stress in human beings (Shaffer, McCraty, & Zerr, 2014) (Thayer et al., 2012) (Prinsloo et al., 2013). The same can be said for an elevated heart rate. While engaged in a fight or flight response, the sympathetic system elevates the heart rate, to effectuate a higher blood flow to skeletal muscles (Fredrikson & Matthews, 1990). This primal reaction makes elevated heart rate an excellent biomarker for stress. The lowering of this rate by an external device would possibly imply that stress is lowered as well.

2.4.2 Skin conductance alteration

Human skin has several functions, varying from protecting vital organs, sensing our outside world, regulating temperature and numerous others (Menon & Kligman, 2009). The conductance of the skin by intra-dermal fluid regulation is of particular interest in this paper. The main reason is that the sympathetic fibers, activated during fight or flight are coupled with palmar and plantar sweat glands and neighboring blood vessels (Lidberg & Wallin, 1981), causes extra fluids to gather in those glands, indicating arousal. Extra fluid transport can be determined by ways of measuring the electrical conductance the skin exhibits, with a better conductance indicating the existence of more fluids. The amount can thus be used in determining the level of activation of the sympathetic nervous system (Miller, 1979).

We can separate two types of skin conductance parameters: Tonic Skin Conductance Levels (SCL) and Phasic Skin Conductance Responses (SCR). SCL is an overall measure of skin conductance, slow in response and can be a proxy for the amount of physical or cognitive activity. SCR is a quick response and more of an indication of sympathetic arousal (Miller, 1979). Measuring elevated SCR as a proxy for sympathetic activity and thus the amount of experienced stress, a feature of the body that will be used as a method in this paper.

2.5 ENTRAINMENT

2.5.1 General – terminology and physiology

Entrainment can be defined as the near perfect synchronization of two oscillating rhythms. In humans, certain temporal biorhythmic processes are able to interlock with another biorhythmic process, thus sharing similar frequencies. It takes place in a manner of natural phenomena, such as a human's circadian rhythm (Golombek & Rosenstein, 2010), heartbeat & respiratory (Seidel & Herzel, 1998) frequencies and musical patterns (Clayton, Sager, & Will, 2005). Different studies have shown that entrainment can take place consciously, such as when people are asked to simulate the beat of a musical piece, or unconsciously, for instance in human communication (Watanabe, Okubo, & Kuroda, 1996).

One of the types best understood is rhythmic sensorimotor entrainment, a variety of entrainment in which people entrain to an outside rhythm with a movement of their own (Thaut, Tian, & Azimi-Sadjadi, 1998). This type is widely utilized by human being in dance, music and sports, but can be generalized to motor movement, with it being an inherent rhythmic process. The human tendency towards rhythmic entrainment can be explained by neural oscillations between sensory and motor processes firing in coordinated bursts, as a natural way of the nervous system's functioning (Wilson & Cook, 2016). Neural oscillations are ubiquitous in the human brain and happen whenever a group of neurons fire in unison or rapid succession (Kotz, Ravignani, & Fitch, 2018). It is argued that the coordination between the neural sensory and motor processes' happens more efficiently when it is done in short synchronized bursts, hereby implying a rhythm that needs to be coordinated in order for it to be effective (Canolty & Knight, 2010). Further evidence supporting this hypothesis was found by Lakatos et al. (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008), where they found that rhythmic sensory stimuli entrain certain sensorimotor neural oscillations causing people to react slower when faced with a reaction task under a rhythmic stimulus.

2.5.2 Entrainment via external stimuli

A significant amount of research has been done in the area of music and its effects on physiological parameters like respiration rate, heart rate (variability) and blood pressure (Etzel, Johnsen, Dickerson, Tranel, & Adolphs, 2006). Later research conducted by Haas et al. (1986) showed a significant correlation between respiratory patterns and the perceived musical tempo, indicating entrainment between the two systems (Haas, Distenfeld, & Axen, 1986). Of particular interest for this paper is the entrainment of cardiovascular processes to external rhythmic stimuli. A general finding is that low complexity and music that is perceived as having calming qualities is reflected in parameters like a lower heart rate, lower blood pressure and higher heart rate variability, sometimes showing at least one physiological parameter to be entrained with music (Tierney & Kraus, 2014). The difficulty with these studies is to isolate one component in the presented stimulus that is dominantly responsible for the alteration of said parameters and whether it is indeed due to entrainment to the stimulus.

Entrainment induced by music was further examined by Bernardi et al. (2006) to include its effect on the cardiovascular system (Bernardi, Porta, & Sleight, 2006). In their study they again selected different types of music, and isolated the tempo component from the music to study the effect it had on heartrate. The results they found were consistent and showed a significant entrainment effect between tempo and heart rate, with heart rate increasing when the tempo rises, a result that is found in other studies as well (Khalfa, Roy, Rainville, Dalla Bella, & Peretz, 2008). Still, the evidence linking only tempo and heart rate by ways of entrainment is thin as of yet, with researchers going as far as to state that a link between these two variables is weak at best (Mütze, Kopiez, & Wolf, 2018) This paper aims to contribute to the better understanding of said effect.

2.6 INTEROCEPTION

Doppel was designed to calm people, with the underlying assumption that the slower heartbeat emulation would entrain the heart to a lower frequency. The idea was that this lower heartrate frequency would then be perceived by the user via interoception, and interpreted as a sign of relative relaxation. However, people are able to influence bodily processes to a varying degree, an ability closely tied to interoception, thus the effects of entrainment might not be noticed equally by all persons.

Interoception shows the degree to which a person is aware of his or her own bodily processes (Wiens, 2005). To some degree, all human begins are aware of their own body, whether it is their overall health, their relative body position or even bowel movement, but the ability to accurately perceive autonomous dynamic processes is quite rare (Craig, 2003). People who are aware of their own bodily processes are better able to appraise their own body reacting to external stimuli, thus helping to regulate emotions to a higher degree than people who are less proficient at interoceptive awareness (Schandry, 1981).

The stress response, as demonstrated in figure 3, is preceded by cognitive appraisal before the body starts to react, with a feedback loop that leads to reappraisal if necessary. Interoceptive sensitivity helps regulate this stress response by constantly reappraising the bodily stress responses (Werner, Duschek, Mattern, & Schandry, 2009).

Measuring interoceptive ability in humans can be done via several methods, with varying degrees of validity. Traditionally, proxies of interoception are taken by ways of tasks to be completed by participants, involving self-reporting on bodily processes during that task. One such task makes use of heartbeat tracking tasks or *cardioception* as a proxy for interoception in general. The original method, widely popularized by Schandry (Schandry, 1981), but originally devised by Dale & Anderson (Dale & Anderson, 1978), has people estimate the amount of their own heart beats that they could detect in 25, 35 and 45 second intervals. Their accuracy is calculated by dividing the self-reported score by their actual heart rate, thus yielding an error rate. The lower the error rate, the better the interoceptive ability is. Other methods include two alternative forced choice trials (Whitehead, Drescher, Heiman, & Blackwell, 1977), where participants are forced to choose between two delayed light flashes after the participants' heartbeat; and multi-intervals tasks, with participants need to align their perceived heartbeats with a button press and an external stimulus (Clemens, 1984), among others.

The measurement of interoception by ways of cardioception has not been standardized and has various results, depending on the method used. Recently, attempts were made in standardizing the theoretical model and its measuring practices, resulting in relatively robust results for the Schandry interoceptive tests (Garfinkel, Seth, Barrett, Suzuki, & Critchley, 2015).

Regardless of the method standardization, interoception is a fairly robust effect found in laboratory experiments. People that have a high degree of interoception are more likely to notice when their body parameters, heartrate among them, change. This assumption holds when physiological parameters are altered due to an external stimulus. Highly interoceptive people notice when their bodies are "tricked" into exhibiting other behavior in the absence of entrainment and can mitigate those changes by appraising the stimulus differently than low interoceptive people. This leads to higher arousal for high interoceptic people as opposed to low interoceptic people when faced with a stressor and an external frequency, due to the fact that they know what their parameters should be. In the case of cardiac entrainment with an external frequency, people with high bodily awareness are expected to be even less aroused, since they pick up the new frequency and appraise it as them being in a relaxed state.

2.7 HYPOTHESES

In this thesis, we attempt to further investigate the effect that the Doppel device has on acute stress. Based on Azevedo et al. (2017) we expect that the participants who have an active Doppel will be less aroused due to entrainment of the heart to the externally applied haptic stimulus, vibrating in a 20% lower frequency than their resting heartrate. This is summarized in the our research questions, as based on those by Azevedo et al. (2017):

- *Does a person have less anxiety induced arousal by usage of the Doppel device as opposed to a person not using the device?*
- *Can the decrease in arousal be explained by ways of cardiac entrainment?*

These initial research questions are detailed further based on the above presented literature analysis, and lead to the following hypotheses:

Hypothesis 1: The usage of a functioning tactile heartbeat simulating device makes people feel less aroused by having a lower skin conductance and higher Heart Rate Variability than when they are not using a functioning tactile heartbeat simulating device;

Hypothesis 2: Entrainment of the heart occurs when people are stimulated by said device;

Hypothesis 3: People that have a high sense of interoception show a lower decrease in arousal to the stimulation of a tactile heartbeat simulating device than people with a low sense of interoception in the absence of entrainment;

3 METHOD

To test the effect of the Doppel on the perceived stress of persons, we conducted an experiment outlined below. The experiment tries to replicate and thus closely resembles the experiment conducted by Azevedo et al. (2017) with a number of questions and procedures added after conclusion of the original experiment.

3.1 DESIGN

The experiment was conducted with an as close as possible resemblance to the original study, meaning a between-subjects design with 2 groups:

- Participants wearing Doppel with tactile stimulus activated (test group);
- Participants wearing Doppel with tactile stimulus deactivated (control group).

Participants were tracked during the experiment for the following parameters: skin conductance, heart rate, heart rate variability and self-perceived anxiety.

After the original study was concluded, two extra elements were added to further investigate the possible outcomes. Participants were presented with a small qualitative questionnaire and a Schandry interoceptive test (Pollatos & Schandry, 2004; Schandry, 1981) was conducted to measure participants' interoceptive awareness.

3.2 PARTICIPANTS

Participants were recruited via the JSF Schouten database at Eindhoven University of Technology. They were excluded if they had disorders associated with the cardiovascular system, since this could influence the heart rate measurements. All participants were briefed before and afterwards and signed an informed consent.

A total of 60 participants were recruited, age 19 – 86, with a median age of 42,5 years $Range_{Age} = 19 - 86, SD_{Age} = 22{,}3\ ;\ M_{Age} = 42{,}5$. Of our total, 23 participants were female, with 37 male. All participants were capable of writing and giving an oral presentation as requested by the experiment conductor.

The amount of 60 participants was selected based on an a priori power analysis. Based on best practice at the department of Human-Technology Interaction at Eindhoven University of Technology we want to have a power of 0.9 for an experiment like we are about to perform. The design of the test was a between group study. In the original study conducted by Azevedo et al. (2017) the effect size was reported to be Cohen's D = 0.54 with two groups of 26 participants. To determine whether an effect size of this magnitude can be detected we performed a sensitivity analysis, indicating that a total of 60 participants were needed per group to be able to detect this effect size. The direction of the tested effect was known, so the analysis was based on a one-tailed t-test.

All participants signed an informed consent form, approved by the ethics committee of Eindhoven University of Technology's Human-Technology Interaction Ethics Board.

3.3 SETTING AND STIMULUS

All experiment sessions were conducted over a two-week period at Eindhoven University of Technology. The first three sessions were performed in the multi-purpose lab at the Human-Technology Interaction laboratories. Due to excessive heat conditions, the experiment environment was changed to an air-conditioned lecture hall, see figures 4 and 5.

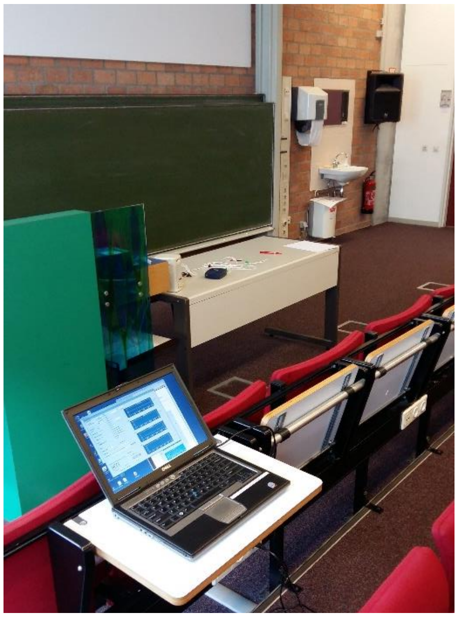


Figure 4: Experiment hall: Experimenter's seat with view on participant's desk

The two locations consisted of a single desk with the sensor material located nearby. On the desk a number of empty sheets of paper and a pen were placed for the participants to use during the experiment.

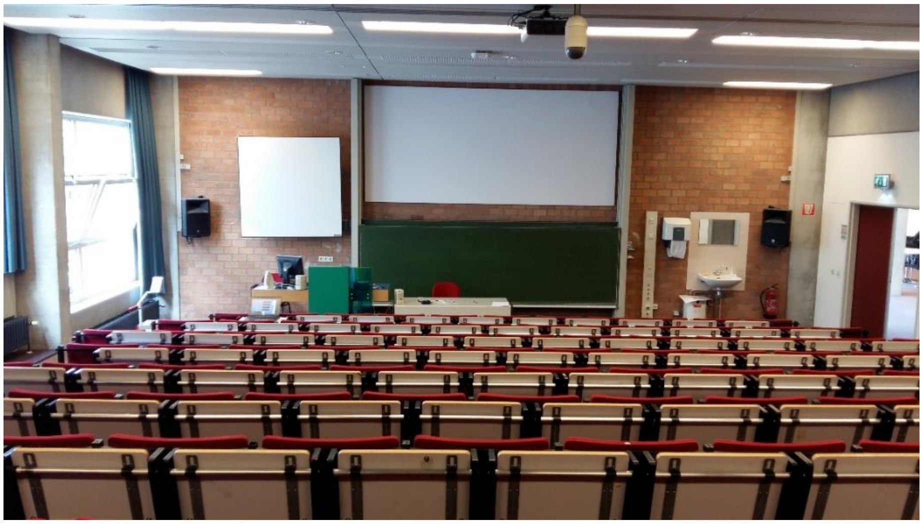


Figure 5: Experiment location

- The multi-purpose lab consisted of a room of about 15 m^2 with a wide window where the experimenter was able to see the participant through a one-directional see-through mirror. Participants were seated facing the window.
- The second major location was a large lecture hall, capable of seating around 90 people. Here participants were seated at a desk normally reserved for lecturers, facing the hall itself.

All participants were fitted with the Doppel device 1.0 version 2016 (www.doppel.com), set to elicit a frequency 20% slower than their resting heart rate for the experimental group. Also, the control group had the watch mounted but switched off. Group sorting was done randomly.

3.4 MEASUREMENTS

Physiological measurements

All physiological measurements were done using a TSMI MOBI-8 recording device with a sample rate of 1024 Hz. Participants were connected to this wireless recorder via two channels: an *ECG-channel* (for measuring HR patterns) with 3 electrodes connected to the upper body, and a *GSR* channel (for measuring skin conductance). The GSR channel was connected via two velcro-mounted dry electrodes to the fingertips of the non-dominant hand, see figure 6. ECG-electrodes were adhesive and placed around the opposing collarbones, with a neutral ground below the lowest rib, see figure 7.

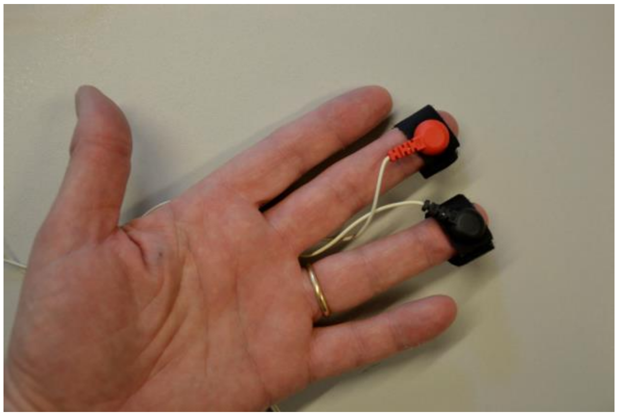


Figure 6: Dry velcro-mounted Skin Conductance electrodes

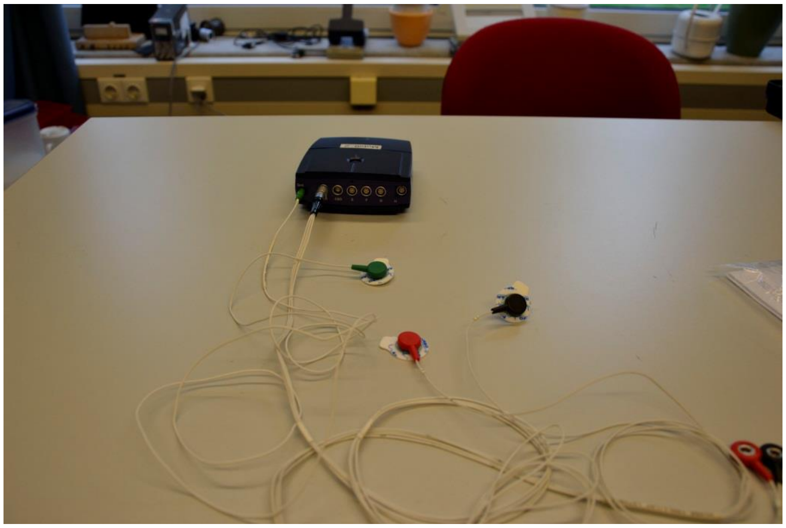


Figure 7: MOBI-device with 3 adhesive ECG-electrodes

Recordings could be digitally marked and stopped/started where necessary during the experiment, in order to separate different stages in the session.

As a final experiment, a short Schandry (Schandry, 1981) test was performed. Participants were instructed to remain seated, with the ECG-electrodes still attached and recording. Instructions were given, along with a form that enabled them to record their findings. The procedure consisted of two simple tasks, performed consecutively. First, in 3 distinct intervals of 23, 32 and 48 seconds, participants were asked to sense their own heartbeats and record them on the form. The experimenter gave a start and stop signal to indicate the interval. The exact time was not conveyed during the procedure. For the second part, participants were asked to record the amount of time that had passed in 25, 35 and 45 seconds, based on nothing but their intuition. This control step is utilized in analysis to control for people that know what their average resting heart rate is and can guess the duration of a second fairly accurately, thus providing an estimated guess and not an interceptive measurement. The utilized form is listed under Appendix A, chapter 6.5.

Subjective measurements

Administration of the State Trait Anxiety Inventory, version Y1 (STAI-1) (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 2010) determined the effect the test had on their subjective stress levels. The STAI-1 questionnaire consists of 20 items on a 4-point scale ranging from: 1 (not at all) tot 4 (very much so), and composed of questions like “I feel frightened, I feel secure” and measures the perceived anxiety of participants. The STAI-Y1 is listed in appendix A, chapter 6.1.

The Brief Fear of Negative Evaluation Scale (bFNE) (Leary, 1983) was administered, to control for the fact that some participants may not find the task sufficiently stressful. The bFNE is listed in appendix A, chapter 6.2

A custom questionnaire added by Azevedo et al.(2017) is composed of 4 items on a 7-point scale and contained questions like: “I like to speak in public”. The questionnaire is not-standardized in peer-reviewed literature and created by Azevedo et al. (2017). The full questionnaire is listed in appendix A, chapter 6.3.

The final custom questionnaire was qualitative and delved deeper into how participants experienced the experiment itself. They were asked to comment on their awareness of the watch, the comfort of the electrodes and the overall intent of the experiment. The custom questionnaire is listed in appendix A, chapter 6.4.

3.5 Procedure

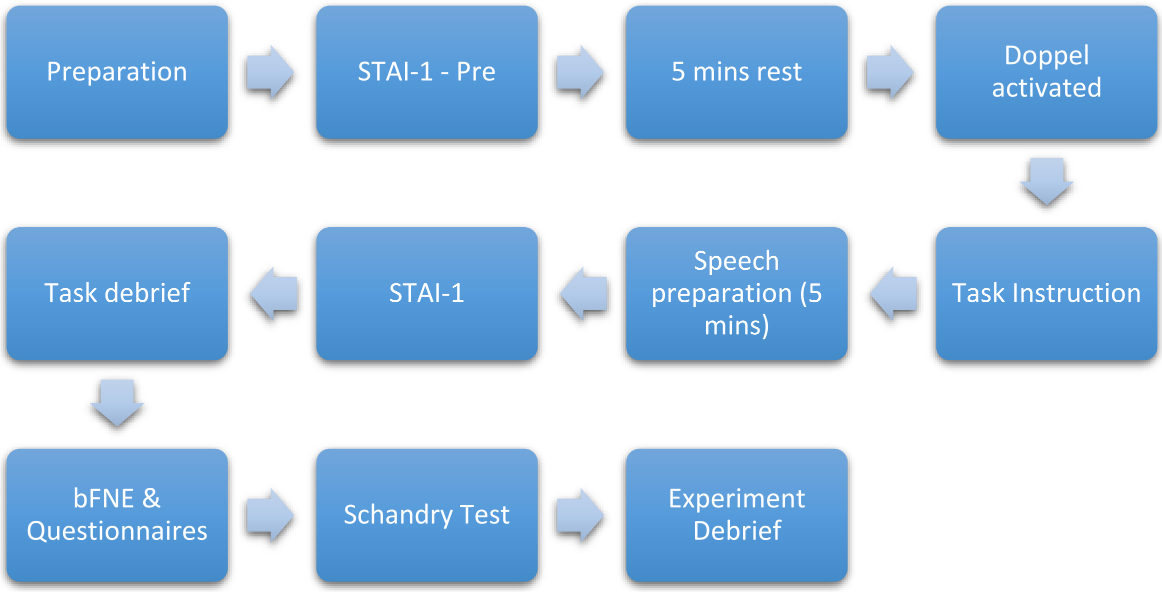


Figure 8: Experiment Flowchart

The overall procedure is listed in figure 8. Overall experiment time clocked in around 45-60 minutes per participant. The different questionnaires were presented to the participants when the appropriate stage of the experiment was reached. This was done In order to combat possible foreknowledge of the experiment's procedure. All questionnaires were filled in by hand.

Preparation

Every participant was invited to test a novel device for measuring blood pressure, to keep them unaware of the actual intent of the study. All of the participants were assigned to a group randomly. The informed consent form, listed in appendix B, was handed out and signed by all participants, with a short introduction and welcome. Their upper body and non-dominant hand were connected to the MOBI-8 and the Doppel-device. Initial testing of correct data transfer was completed after electrode appliance. Paper and pencil were provided to the individuals upon arrival.

Pre-task STAI-1 administration & baseline measurement

Before instructions were handed out about the stressful task, participants were required to fill in the pre-experiment STAI-1. The form was collected shortly afterwards. Participants were left alone to rest for approximately 5 minutes. The resting heart rate and skin conductance was recorded as a baseline measurement and the last couple of samples of ECG were used in the subsequent procedure.

Fitting the Doppel watch

Participants were required to wear a watch and were told that this particular device would measure their blood pressure. In reality, the watch was capable of eliciting a vibrating, tactile stimulus (to which half of the participants were subjected, depending on the condition). Before the commencement of further explanation, the experimenter explained that the device would be switched on, to start the blood pressure measurement and not to be alarmed if they felt any vibration upon booting of the device. The device was then switched on and switched off after about 10 seconds for the control group (without telling them) and kept on for the experimental group. The frequency of the tactile stimulus itself was based on the previously recorded personal resting heart rate of said participant($MeanHR = 76, SD = 11.6$), modified to be 20% slower($MeanHR = 60, SD = 9.2$).

Stressful task

An instruction was given that participants needed to give a short presentation of approximately 5 minutes, for which they were given 5 minutes to prepare. Contents of the presentation was up to them, however the subject needed to be animal usage for research purposes. Pen and paper was permitted in the preparation. The researcher then left the room for 5 minutes.

Post-STAI, bFNE and open Question questionnaire

After 5 minutes, the experimenter returned and participants were presented with a second STAI-1 form and were requested to complete it once more. When finished, they were informed that they did not need to perform the presentation. The bFNE, that measures public speaking anxiety, was then presented and filled in. The bFNE was only done once, after they were told that they did not need to present. This was done to circumvent any pre-presentation anxiety that may have skewed the results. Finally, the two custom questionnaires were given to complete at their own leisure.

Schandry task

Participants were finally subjected to a Schandry test (Schandry, 1981) with a time keeping test at the end of the procedure, as described in section 3.4, Their ECG's were recorded during the task.

After ending the test, participants were decoupled from the MOBI-device, debriefed and compensated accordingly.

3.6 ANALYSIS

Repeated measures tests for the psychophysiological tests were conducted with time as within group factor and condition as between group factor. All of the recorded samples were marked in time digitally to signify different stages in the experiment.

3.6.1 BACH AUC pre-processing

We have pre-processed the signal for SC- AUC measurements according to Bach et al. (Bach, Friston, & Dolan, 2010), which was the method used in the experiment according to Azevedo et al. The signal was bandpass filtered (0.015 – 5 Hz) using a 2nd order Chebyshev filter. Afterwards the average is calculated by determining the lowest value of the signal and subtracting that from every sample, and subsequently integrating by time/amount of samples. This creates a residue signal that cuts off the intercept and solely considers the peaks above said intercept.

3.6.2 Psychophysiological measurements

Skin Conductance

One participant was excluded from this test due to a measurement error. The main dependent variable in this experiment is the skin conductance, measured in microSiemens (μS), where a higher value indicates a higher conductivity, which equals a higher sweat production and thus more arousal (Dawson, Schell, & Filion, 2016). The samples were pre-processed using Matlab 2018b, utilizing custom code available from the experimenters. For pre-processing purposes, a Chebyshev 2nd order filter, with a bandpass and stopband cut-off frequency of 0.015 Hz – 5Hz was applied.

Finally, the signal was aggregated in two different ways to see how the participants responded to the proposed tests. First, following Azevedo et al. (2017), and based on the method presented by Bach et al (Bach et al., 2010), we measured the Area under the Curve (AUC) based on the following equation:

Equation 1: AUC - Skin Conductance calculation

$$SC_{AUC} = \frac{1}{T} \int SC(t)dt - SCL$$

AUC = Area Under the Curve (as defined by (Bach et al., 2010)

SCL = Minimum skin conductance level during integration period;

t= time in seconds.

T = time interval in seconds (5 minutes for baseline and speech preparation);

The second aggregation was conducted by following a more traditional SC analysis by ways of the following equation:

Equation 2: Skin Conductance Average

$$SC_{avg} = \frac{1}{N}\sum SC(t)$$

All resulting variables were tested for normality, results of which are found in the following chapter.

Heartrate

Heart rate analysis was carried out by two programs: RR-detect, a custom piece of software provided by the HTI-labs, and HRVAS v1.0.2, a Matlab 2018b program for advanced heart rate analysis. All of the programs are available from the author (Boschman, 2012).

The RR-detect program analyzes the QRS-complexes and generates inter-beat intervals (IBI) values for further analysis, after possible exclusion of said IBI’s. An inter-beat interval is the time between two consecutive R-peaks in het ECG, with The R-peak being the highest measurable cardiac signal in a standard ECG. See figure 9 for further reference.

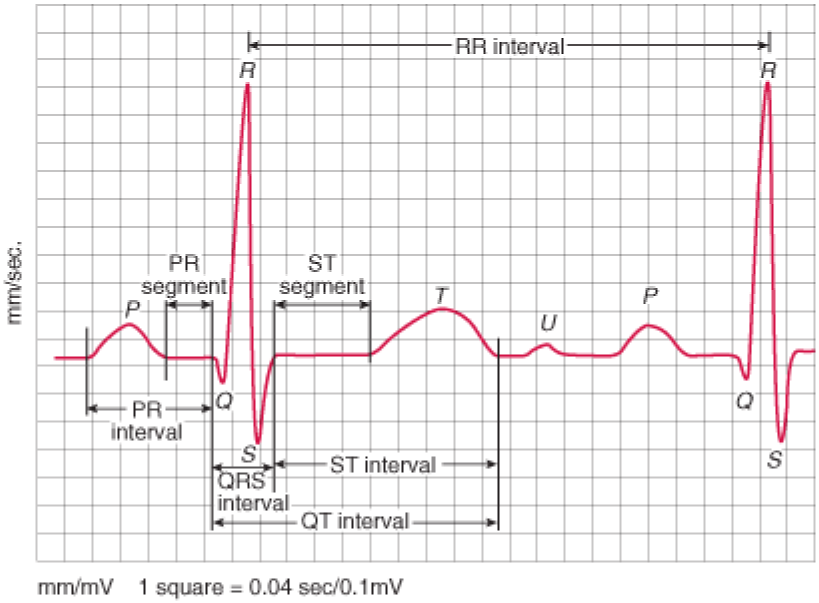


Figure 9: R-R interval in a standard ECG

The program was set to reject IBI’s smaller than 0.4 sec and IBI’s that exceeded 1.4 sec, due to possible measurement errors that could result in these extremes. Furthermore, whenever the considered IBI exceeded 0.5 or higher than 1.5 times the sliding average over the last XX values, the IBI was excluded as well.

HRVAS analyzes the IBI's and generates statistics based on them. The generated statistics in the baseline measurement, pre-task and post-task were used to determine time-domain measurements like average heartrate (HR), heartrate descriptive statistics and heartrate variability (HRV) using the Root Mean Square of Successive Differences (RMSSD) of the heartbeats (Allen, 2002).

The utilized RMSSD is calculated in the following manner:

Equation 3: RMSSD calculation

$$RMSSD = sqrt\,\{mean((RR_{i+1} - RR_i)^2)\}$$

3.6.3 Statistical Assumptions testing & analysis

All of the quantitative analyses of the physiological data were done by ways of repeated measures ANOVA with *time* as a within group factor and *condition* as a between group factor. This is applicable to the following physiological parameters:

- Skin Conductance Averages
- Skin Conductances AUC-corrected;
- Heart Rate Variability – RMSSD;
- Heart Rate;

The variables in the between-measures ANOVA are checked for the following assumptions:

- Each group sample is drawn from a normally distributed population;
- All populations have a common variance;
- Within each sample, the observations are sampled randomly and independently of each other;

Quantitative psychological data was analyzed using conventional t-tests. This is applicable to the following tests:

- STAI-Y1;
- bFNE;
- Custom Azevedo Questionnaires;

For the t-tests to be valid, we checked if the following assumptions were met:

1. Continuous variable;
2. Two groups:
3. Normal distribution
4. Homogeneity of variances:

All the variables were analyzed and subsequently transformed using Tukey's ladder of powers. The transformations, where applicable, were transformed in a similar manner for variables in the same analysis. The specific transformation was mentioned per analysis.

3.6.4 Subjective measurements

Four quantitative questionnaires were administered, with the STAI-Y1 being administered twice. The STAI was analyzed using the primer provided by the author (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). Here the negative items were reversed and all items aggregated and subsequently analyzed with t-tests after checks for normality. A higher score indicates higher self-reported anxiety.

bFNE analysis was performed according to the revised primer by the author (Carleton et al., 2006), aggregating the scores and analyzing them with t-tests after normality checks. A higher score indicates higher anxiety due to public speaking.

The interoceptive data of resting heartrate and reported heartrates, measured while conducting the test, are transformed to a full minute. With the difference between reported and measured beats subtracted to achieve a final score. This procedure is done 3 times in intervals of 25, 35 and 45 seconds, of which the average will be taken. The subsequent counting test is transformed in the same manner.

A heartbeat perception score was calculated according to the following equation, summing over the 3 intervals:

Equation 4: Schandry Perception Score

$$Perception\ score = 1/3 \sum \left(1 - \frac{|\#recorded\ heartbeats - \#counted\ heartbeats|}{\#recorded\ heartbeats}\right)$$

High scores (with a maximum of 1) indicate absolute accurate heartbeat perception ability. A cutoff score of .85 was used to categorize subjects either as good or poor heartbeat perceivers. This is in accordance with the cutoff selected by Montoya et al. (1993) and Weitkunat and Schandry (1995).

The Time Awareness scores were calculated according to the following equation:

Equation 5: Time Perception Score

$$Time\ awareness\ score = 1/3 \sum (|Reported\ seconds - actual\ seconds|)$$

The lower the deviation, the better the time awareness. This metric will be used to control for possible "educated guessing" by participants with knowledge of their resting BPM and good time awareness.

3.6.5 Qualitative analysis

All interviews were transcribed and digitalized, labeled per question with a content analysis. Different groupings were aggregated based on their answers and condition.

3.6.6 Experiment continuation analysis

Since this is a replication, the original effect size based on the arousal level was reported as Hedge's $g = 0.54$ with two groups of 26 participants. Because the direction of the effects is known, based on the previous experiment, a one-tailed analysis was conducted. The power analysis showed that in order to attain this power, the experiment would have to have two groups of 60 participants. In order to make the experiment feasible and less costly, an alternative was proposed involving sequential analysis at the moment half of the participants are run.

For sequential analysis, WinLD v2.0 was utilized, a Python based statistical analysis program. The bounds for a possible continuation of the study are based on sequential analysis using on the Pocock-method (Pocock, 1977), one-sided, using 2 equally spaced intervals due to the fact that our experiment continuation analysis stops at exactly half of the required participants. The Pocock-method was selected due to the uncertainty of the actual effect size. Based on this method, we consider the study to be worth continuing when:

A) The likely effect size exceeds 0.3;
B) The p-value of the repeated measures ANOVA on the skin conductance response > 0.031;

In this study, we have collected data for a total of 60 participants, 30 in each group. Based on the findings, a recommendation will be made later in this report to either halt collection of data or continue data collection if possible in the future. If the above described analysis attains a $p < 0.031$, the study will be terminated early due to it being successful. If the study approaches $p = 1.0$, the study will be terminated due to the unlikeliness of it reaching significance.

4 RESULTS

In this chapter we will be discussing the results of this study. We aimed to investigate psychological and physiological parameters measuring the amount of arousal participants experience when faced with a stressful task, while wearing a tactile, heartbeat simulating device.

In paragraph 4.1 we present arousal levels based on skin conductance levels, along with the heart rate and heart rate variability during the stressful task. The results of the self-reported arousal levels measured by the bFNE and STAY-1 questionnaires, are presented in section 4.2. The Schandry interoception test is covered in section 4.2.4. Finally, the results of the custom questionnaires are presented in section 4.3.

Due to the nature of the experiment, all samples listed in this study are drawn independently, the two groups are categorical and the dependent variable continuous.

4.1 PHYSIOLOGICAL MEASUREMENTS

4.1.1 Skin conductance Averages

The calculated averages of the skin conductance are aggregated according to equation 2, compared and analyzed using a repeated measures ANOVA, with *time* as a within group measure (the difference between the start and end of the stress inducing task) and *condition* as a between group factor (to differentiate between the two *test* and *control* conditions) with *measured skin conductance average* as the dependent variable.

Two participants were excluded from the skin conductance results due to technical difficulties resulting in measurement errors, both from the control group. Normality was rejected for both variables, therefor the skin conductance variable was transformed using a square root transformation, which made the control variable normal. Upon inspection of the histogram, the distribution was deemed sufficiently normal to go ahead with the repeated measures ANOVA. The test showed a statistically significant effect for the *Time* measure ($F(1,56) = 158.599$, $p<0.001$), making the task sufficiently stressful. However, the Time*Condition measure showed a non-significant result ($F(1,56) = 3.873$, $p = 0.54$). The Condition-effect was non-significant. Effect size $f = 0.26$ for the interaction effect Time*Condition.

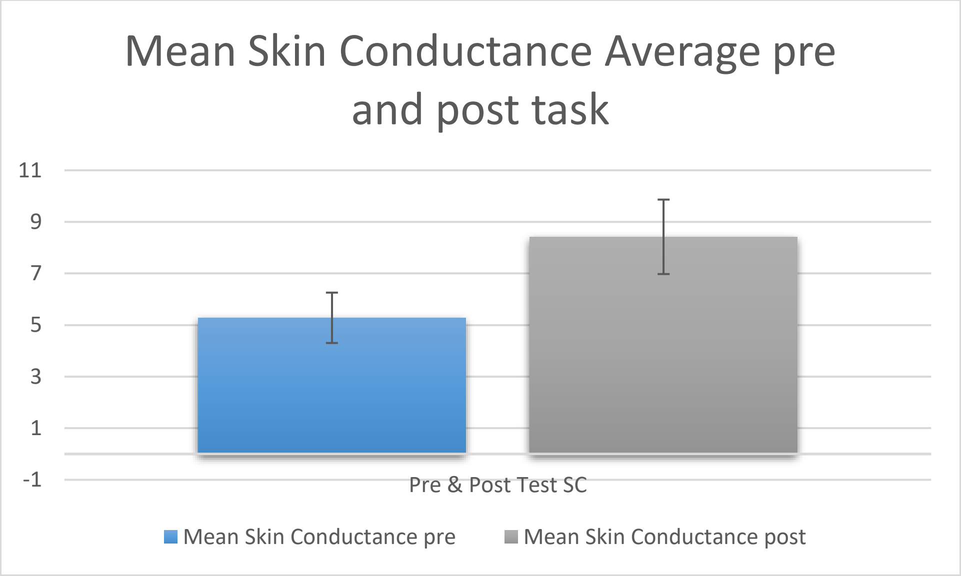


Figure 10: Mean Average overall Skin conductance pre and post task with 95% CI

The differences between the two conditions in skin conductance is shown in figure 10.

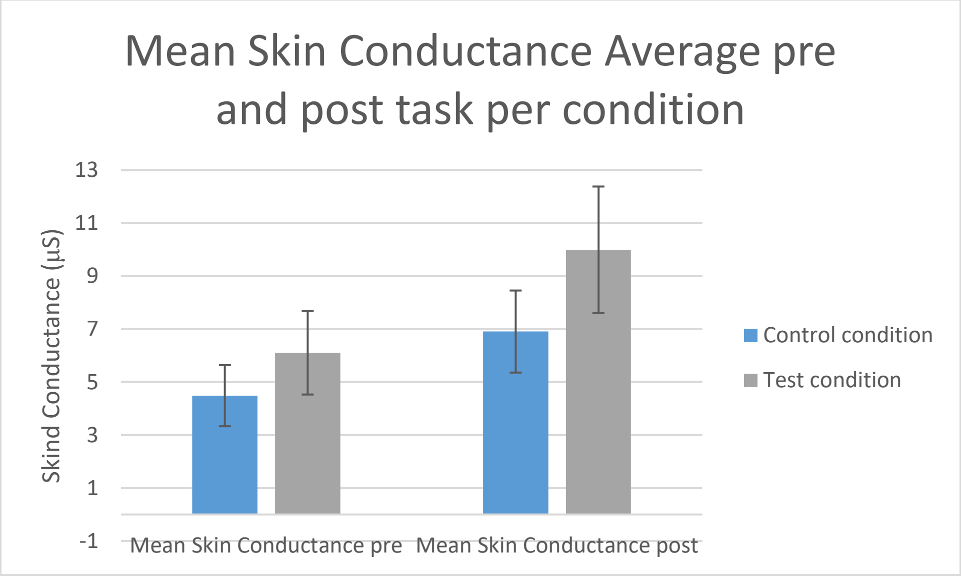


Figure 11: Mean Average skin conductance pre- and post-task per condition with 95% CI

4.1.2 AUC measurements

Two participants were excluded from the data due to a measurement error. The obtained variables are tested for normality, which was rejected for both conditions. Upon closer inspection of the data, five outliers were removed and the square root of the variables were taken. After this transformation, the variables are considered normal. Similar to the average value in skin conductance, we analyzed the results using equation 1 and a repeated measures ANOVA with *Time* as a within group factor and *Condition* as a between group factor. The SC-AUC was taken as a dependent variable.

Same as the general average in the previous paragraph, the effect TIME has on participants across conditions is significant (*F*(1,51) = 14,657, *p*<0.001), once again suggesting the task is properly stress-inducing. The interaction effect TIME*CONDITION is once again non-significant (*F*(1,51) = 1,171, *p*=0.284). Main effect of Condition is non-significant. Effect size is *f* = 0.14 for the interaction effect.

The observed effect in the AUC analysis, though not significant yet, is pointing to the other direction, as shown by figure 11.

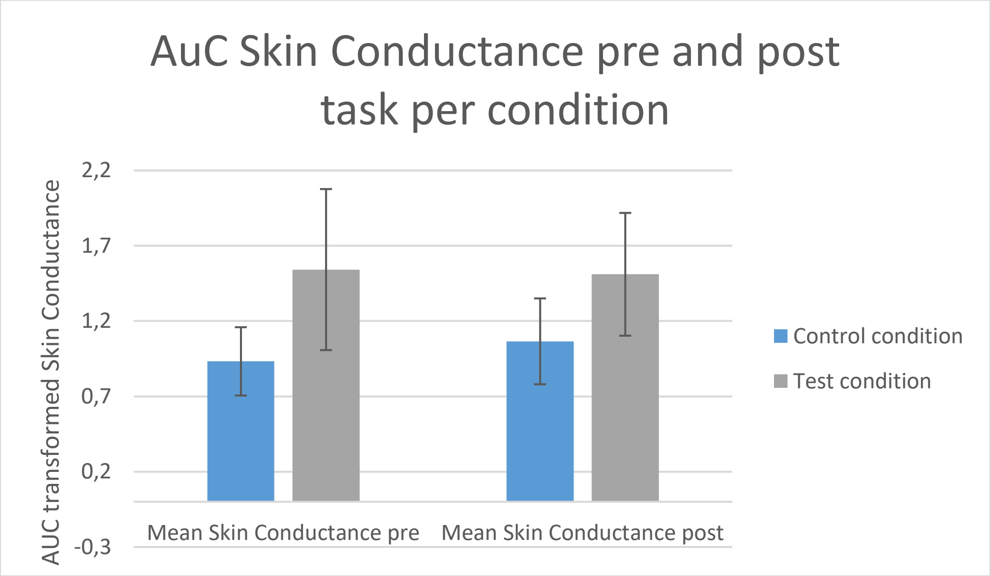


Figure 12:AuC-mean skin conductance pre- and post-task per condition with 95% CI

4.1.3 HRV

Two measurements were excluded due to a measurement error. The variables for heart rate variability, using the RMSSD as a metric were tested for normality, all rejected. A Tukey Ladder and box-cox test yielded no results, even after examining outliers. A repeated measures ANOVA was conducted, acknowledging that interpretation is more difficult. The two-way repeated measures ANOVA was run with *Time* as a within group factor and *condition* as a between group factor, with RMSSD as the dependent variable.

The TIME condition (*F*(1,56) = 0.305, *p* = 0.583) showed a non-significant result, hinting that there is no significantly lower or higher heart rate variability before or after the task. The Condition effect was non-significant. The interaction effect TIME*Condition (*F*(1,56) = 0.290, *p* = 0.261) also indicated no significant effect between the two variables and conditions. The effect was in the hypothesized direction, and could reach significance when study is completed with the full roster of participants.

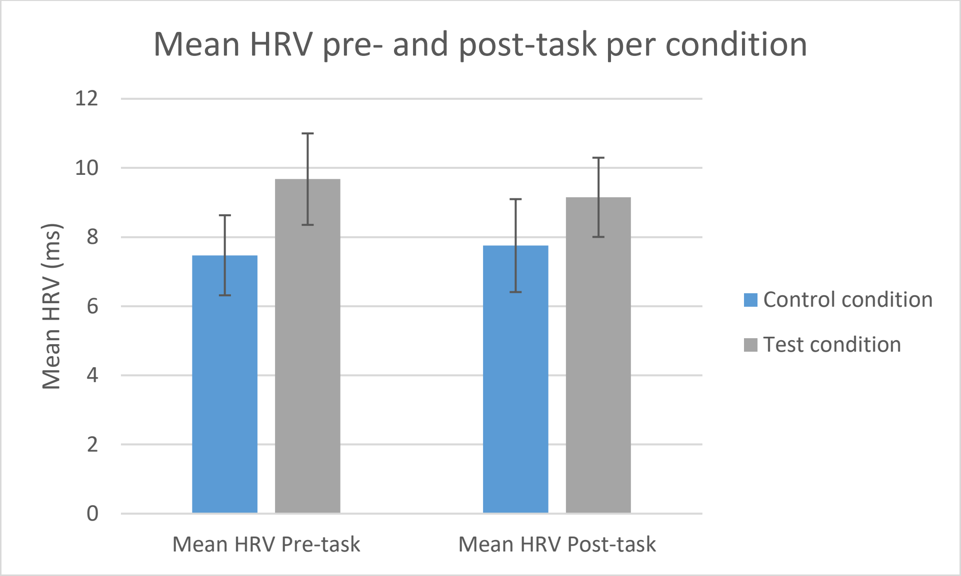


Figure 13: Mean HRV pre- and post-task with 95% Confidence Interval

4.1.4 Entrainment

We excluded two participants due to a measurement error. Subsequently, we pre-processed the HR variable in the two conditions due to non-normality by excluding two outliers, resulting in both variables to be considered normal. A two-way repeated measures ANOVA was conducted to analyze the entrainment taking place during task preparation, expanding on our hypothesis that possible relaxation is due to entrainment between the haptic device and the heart. The ANOVA was run with *Time* as a within group factor and *condition* as a between group factor on the dependent variable HR. We expect that cardiac entrainment takes place and results in a significantly lower heart rate in the *test* condition.

The test showed a significant effect in the TIME condition. (*F*(1,54) = 72,613, *p*<0.001), indicating a significant heart rate elevation during task preparation. The interaction effect TIME*CONDITION once again showed a non-significant effect in the expected direction (*F*(2,54) = 0,86, *p* = 0.36). The was no significant main effect of CONDITION. Effect size *f* = 0.12 for the interaction effect Time*Condition.

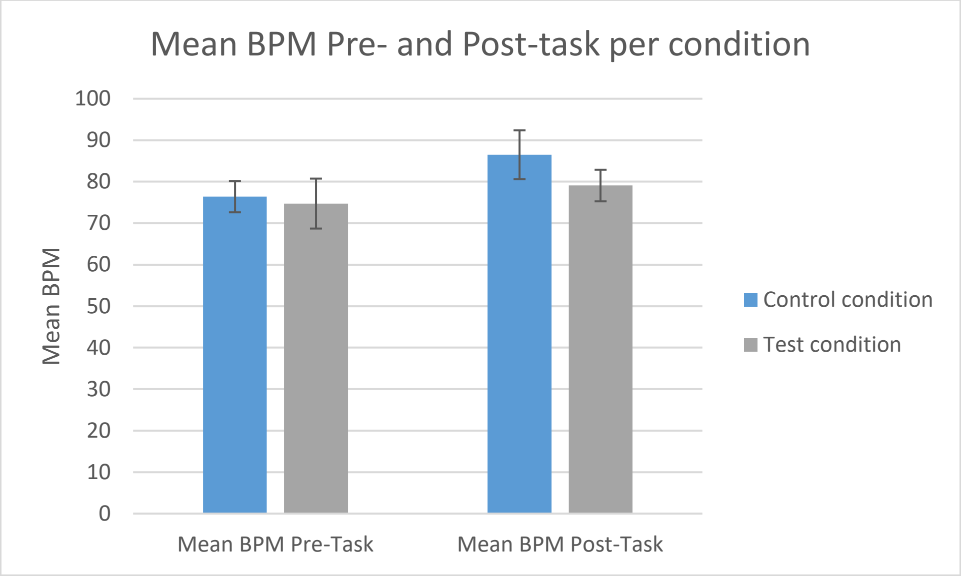


Figure 14: Mean BPM pre- and post-task with 95% Confidence Interval

4.2 PSYCHOLOGICAL MEASUREMENTS

4.2.1 Self-reported anxiety (STAI)

No participants were excluded from this analysis. Analysis revealed that the pre- and post-STAI variables are not considered normally distributed. Therefore, a log transformation was performed, which rendered both variables normal.

A two-way repeated measures ANOVA was conducted, with t*ime* as a within group factor and *condition* as a between group factor, using the STAI-Y1 scores as a dependent variable. Based on this ANOVA, a statistically significant effect was found for the *time* factor ($F(1,58) = 27,49$, $p < 0.001$), suggesting that people experienced the task as a stressful event. The interaction effect time*condition was non-significant, with the effect pointing in the expected direction ($F(1,58) = 0,6$, $p = 0.442$). The effect size $f = 0.1$.This effect could reach significance when the full amount of participants is used in a follow-up study.

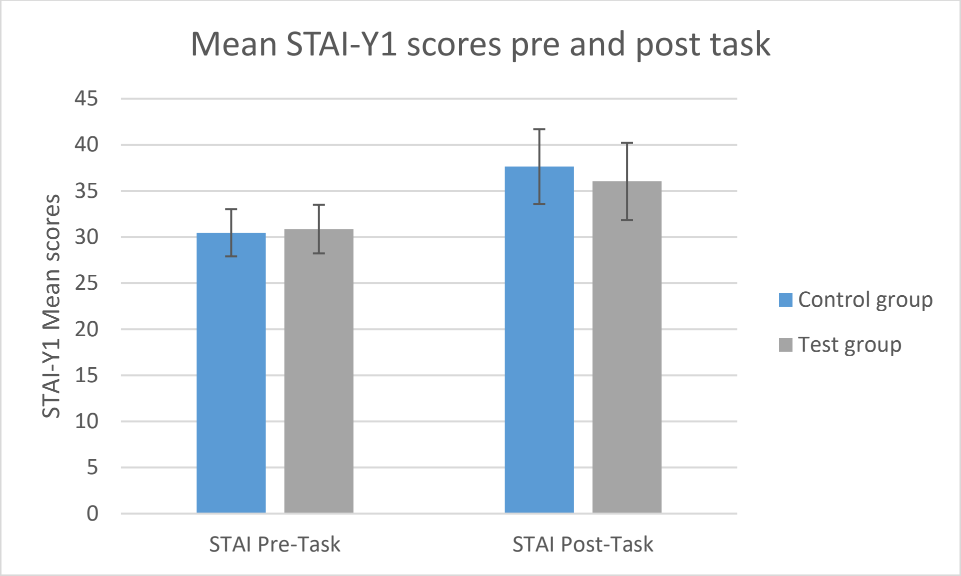


Figure 15: Mean STAI-Y1 scores pre- and post-task with 95% Confidence Interval

4.2.2 Self-reported Speech Anticipation Attitude (bFNE)

The variables were tested to be non-normal. Subsequently, the variables were log-transformed to have normality non-rejected. Due to the fact that this questionnaire was administered once after the task preparation was completed, an independent t-test was performed to examine the difference between the two conditions, with the aggregated bFNE-scores as a dependent variable.

Results show that the control group (M = 1,45, SD = 0.169) does not differ significantly from the test group (M = 1.409, SD = 0.160), *t*(58) = 0.965, *p* = 0.339, Cohen's *d* = 0.79. This result indicates that the two groups have roughly the same attitudes towards public speaking.

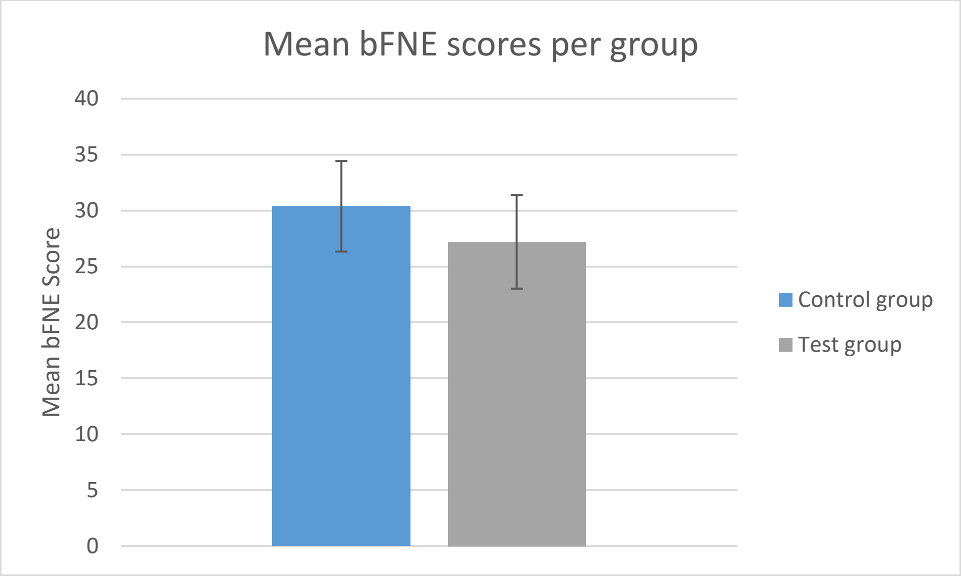


Figure 16: bFNE Mean scores per Group with 95% Confidence Interval

4.2.3 Azevedo custom questionnaires

The custom questionnaire administered by Azevedo consisted of four separate questions, each analyzed with independent t-test. All analyses were done on an item-by-item basis to analyze the difference between the two groups per question. Normality of all the variables was rejected, with three of the questions transformed using its square root and the Social Anxiety question log transformed.

The difference between the control group (M = 3.84, SD = 1.9) and the test group (M = 4.58, SD = 2.29) for the question *“I like to speak in public”*, was non-significant (*t*(58) = -,1,15, *p* = 0.253), Cohen’s *d* = 0.35. As for the question *“I can easily become anxious in social situations”* the difference between the control- (M = 4.29, SD = 2.00) and test group (M = 3.82, SD = 1.81) was non-significant (*t*(58) = 0.515, *p* = 0.608), Cohen’s *d* = 0.24, with the test group having slightly lower scores overall. This indicates an effect in the hypothesized direction, which could reach significance in a follow-up study. The question *“My experience during the task preparation was:*”, was posed twice. The first instance examined the degree of experienced stress, where the difference between the control group (M = 4.77, SD = 1.61) and the test group (M = 4.82, SD = 1.77), Cohen’s *d* = 0.03,was once again non-significant (*t*(58) = -0,058, *p* = 0,954). The second instance examined the degree of pleasantness during the task, where the difference between the control group (M = 5.03, SD = 1.52) and the test group (M = 4.96, SD = 1.59), Cohen's *d* = 0.04, was non-significant as well (t(58) = -0,037, p = 0,97).

Overall, the two groups did not differ enough in their responses to help explain group differences.

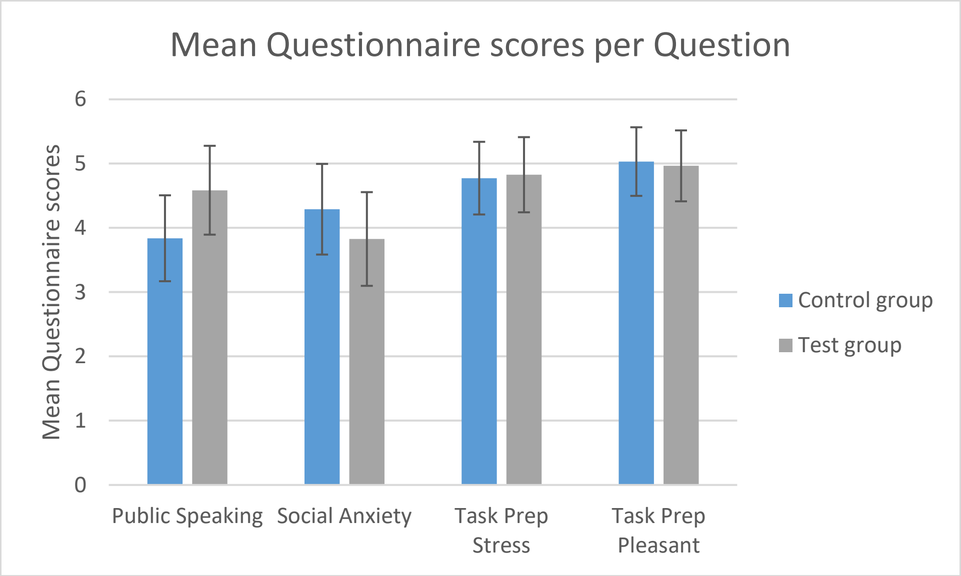


Figure 17: Mean Azevedo Custom Questionnaire Scores with 95% Confidence Interval

4.2.4 Interoceptive measurements (Schandry)

Due to a measurement error, 3 participants were excluded from further analysis. The two variables have normality non-rejected. A linear regression was performed to see the connection between reported and actual heartbeats, taking the averages of the time keeping results and the average of the Schandry task-results. An $R^2 = 0.247$ indicates a weak link between actual and reported heartbeats. The regression's F-statistic does show a significant difference between the two scores: F(1,56) = 18.017, p < 0.001. In figure 18 and 19, the mean scores for the population is shown for the Schandry task and the time-keeping task, respectively.

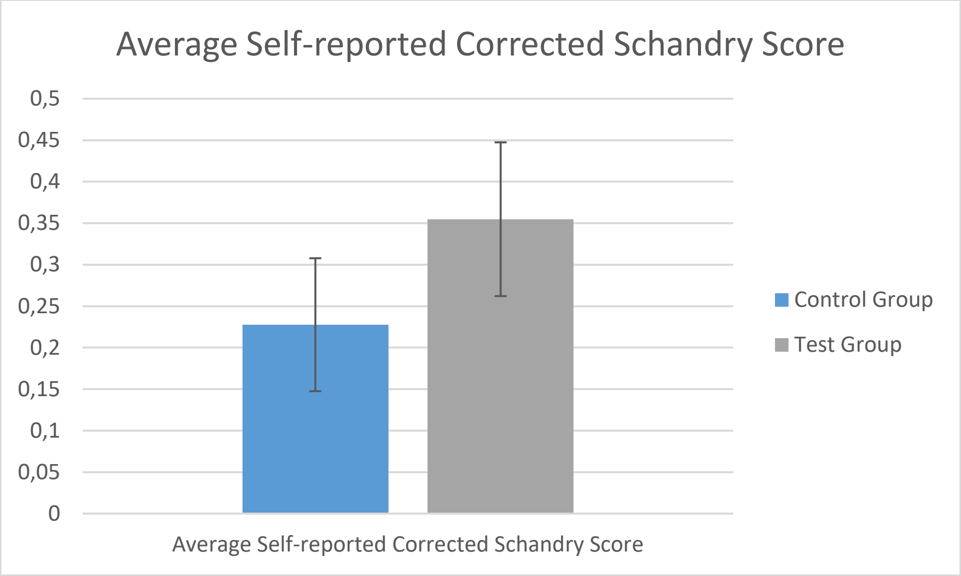


Figure 18: Average Self-Reported Corrected Schandry Score with 95% CI

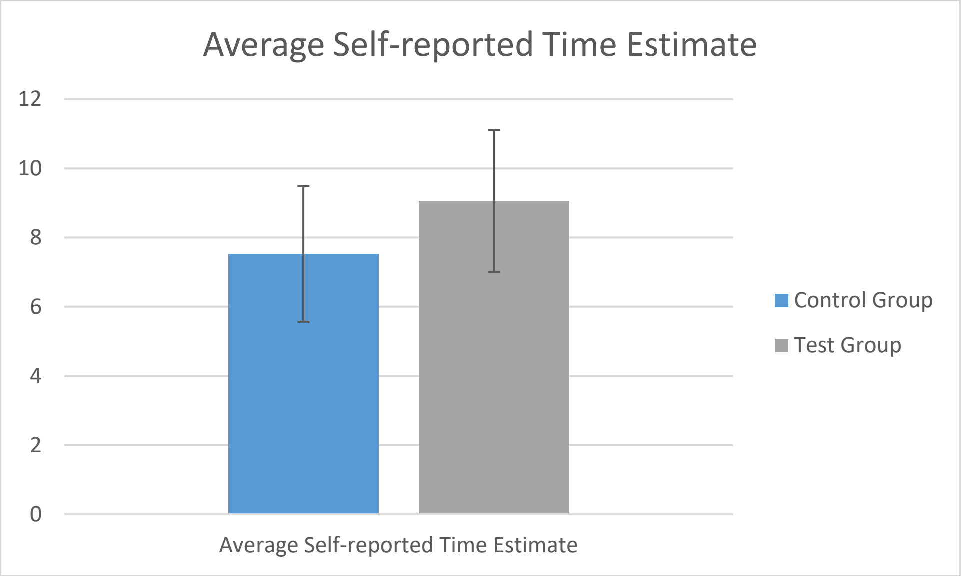


Figure 19: Average Self-reported Time Estimate with 95% CI

According to Pollatos et al. (Pollatos, Kirsch, & Schandry, 2005), an interoceptive awareness score of 0.85 constitutes a good heartbeat perception. In our research, not a single participant scored higher than the 0.85 error margin threshold, labeling all participants as less than highly interoceptive. The interoceptive awareness scores for the full sample were significantly lower than the previously mentioned 0.85(Min = 0, Max = 0.71, M = 0.36, SD = 0.23).

Due to not a single participant breaching the 0.85 threshold, a subsample was examined to further explore the findings, ultimately aiding us in answering hypothesis #3. We selected the cases that were more than 1 SD above the mean, thus having a heartbeat interoceptive score of approximately 0.6. This cut the sample down to 9 cases with a heartbeat interoceptive score of 0.6 or higher. Based on equation 5 in this study, a low time counting score indicates a high risk of counting heartbeats, making the Schandry results less reliable. Of this subsample, 3 cases had a time counting score of 4 or lower, increasing the risk that this subgroup "counted" their heartbeats, skewing the result. To mitigate this risk, they were excluded from this subsample, leaving us with 6 cases, all with a time counting score of 7 or higher. Two of the cases were in the control group, 4 of them in the test group. It should be noted that this sample is very small, making interpretation very hard, if not impossible. Regardless, the results of high interoception on skin conductance for the average and AUC signals in this high interoceptive subsample are shown below.

For both of the dependent variables, normality was non-rejected in both cases, as well as equality of variances. Two different repeated measures ANOVA were conducted, taking the Average Skin Conductance Response, and the AUC Skin Conductance Response as dependent variables, both with TIME as a within group factor and condition as a between group factor.

In the Average test, the Time factor was significant (F(1,5) = 95.1, p <0.001). The interaction effect TIME*Condition was significant as well (F(1,5) = 17.1, P = 0.014), suggesting that in this subsample, the difference between the two conditions over time was significant, something that was unexpected given the hypothesis.

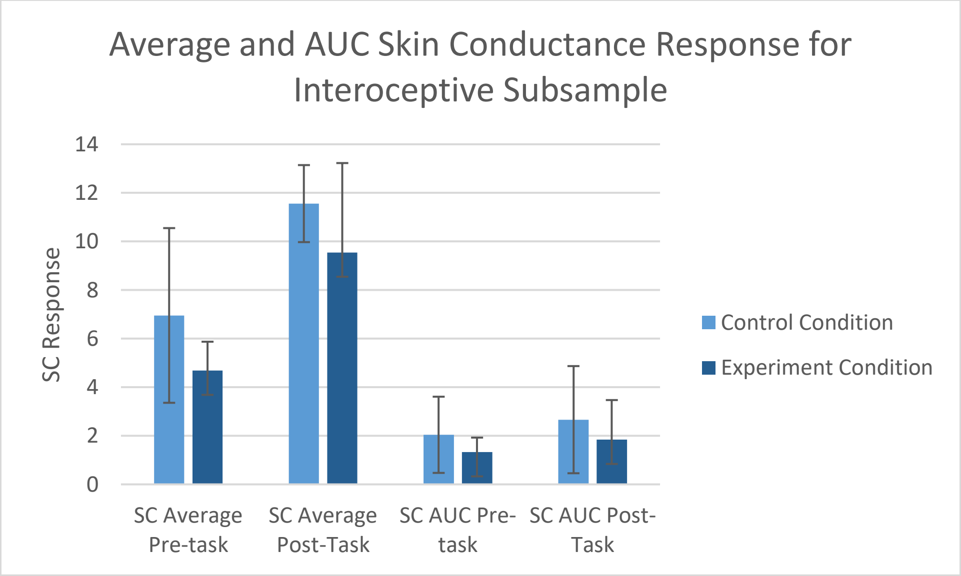


Figure 20: Average and AUC signal Skin Conductance Response with 95% CI for Higly Interoceptive Subsample

In the AUC test, the measure TIME was, surprisingly enough, non-significant (F(1,5) = 6.11, p = 0.069), with the interaction effect Time*Condition non-significant as well (F(1,5) = 6.22, P =0.067. The interaction effect being non-significant is interesting, would it not be that the between conditions effect TIME is non-significant for both conditions in the group.

4.3 CUSTOM QUESTIONNAIRES

The custom questionnaires were administered after completion of the task preparation and consisted of 12 open items to be filled in by the participants. All participants managed to fill in the questionnaires fully.

A common theme in the questionnaire was that in general the participants believed they needed to give the presentation, with 80% responding positive to this question. Moreover, about half of the participants were in one way or another uncomfortable with presenting in front of a live audience. Some of them were even aware of the possibility that they may be tricked due to extensive experience with psychophysical experiment, but still reported to be anxious about presenting: "*Having knowledge about these kind of experiments, [I did] not really [believe I needed to give the speech]. Even so, [I believed it] a little bit, because the possibility is always there*".

Of the reported cases, 30% was aware of the Doppel during the experiment in varying degrees, ranging from "*a little bit, but if my normal watch wasn't broken this month, I probably wouldn't have",* to *"yes, it was vibrating a lot in a steady pace".* Of those 18 cases, 7 indicated that it influenced their responses in the test, with 3 people indicating that even though they did not notice the watch, it still influenced their reaction in the experiment.

In 27% of the cases, the participants felt calmed by the vibrating device, with half of the cases being in the control condition, indicating that some of the participants felt calmed by just the presence of the device, as a participant indicated: "*A bit, at least the vibrations weren't uncomfortable"*. Other participants in the test condition indicated that the vibrations were actually more of a stressor than a relaxing stimulus. For instance, one response was: "*To be honest, It increased [me feeling tense]. So I'm not really calm".* Others noted that the vibrations actively disturbed them: "*No [I did not feel calmed]. [The watch] kept on disturbing me whenever I tried to put my mind in a pleasant mood".* Some people tried to stay calm, but the watch made that endeavor harder, as stated by one participant: *"I tried to [actively] stay calm, but the watch kept on interrupting me".*

The devices needed to measure the participant's reaction (ECG electrodes, Doppel and hand-mounted sensors) were deemed to be comfortable in 93% of the cases, with 7% noting that the watch and/or the hand-mounted electrodes were too tight. In one case, the electrodes gave the sense that he/she should not move the hand, thus it being a bit distracting.

About 83% indicated that they would be comfortable wearing the Doppel regularly and even for an entire day, distributed evenly among the control and test conditions.

The actual intent of the Doppel device was unknown to 80% of the participants, with 20% having varying degrees of suspicion about the actual function, none of which were correct about the actual function. 7% stated that it influenced the way they conducted the test: two of them indicated that it made them more calm, the other two conveyed that it made them a bit more stressed.

5 DISCUSSION

The presented research set out to examine the effect that a haptic, heartbeat emulating device has on the measured and perceived stress on humans. The procedure focused on replicating the original study by Azevedo et al. (2017) as precise as possible, with added elements to further examine the underlying causes for possible deviation from the original research's findings. The expectation was, in line with the findings of Azevedo et al.(2017) and the underlying research, that a subtle vibration at 20% below their normal heart rate would lower participants' arousal and perceived stress, thus calming them. We also hypothesized that this effect was due to cardiac entrainment, resulting in a significantly lower heartbeat when presented with the stimulus. However, neither of the expected effects were found.

The last hypothesis focused on the degree to which participants were aware of their bodily state by showing a high degree of interoception, and whether that influenced their arousal levels during the procedure.

5.1 MANIPULATION CHECK

5.1.1 Task effectiveness

As the basis for the experiment, the speech preparation task needed to be sufficiently stressful for us to be able to detect differences between the groups. Across all measures taken: Skin Conductance Response, Skin Conductance Response AUC, Heart Rate, Heart Rate Variability and subjective responses we measured significantly higher values, indicating that the task preparation worked adequately as a stressor.

5.1.2 Self-reported measures

Results show that there was no difference across the conditions in the attitudes of people towards public speaking. The data indicates that participants from both conditions are equally likely to find a public speaking task stressful or non-stressful. In the self-reported measures taken, we did see a significant difference in the self-reported stress response, indicating that the task preparation was a stressful experience. There was no significant difference between both conditions after the task preparation, making the Doppel's effect marginal in mitigating subjective stress response.

In line with the article Azevedo et al. (2017) published, we conducted a simple custom survey next to the verified bFNE en STAI-Y1, asking for people's attitudes towards public speaking, their social anxiety and the task's stressfulness and pleasantness. In our sample, no significant differences were found, once more indicating the homogeneity in both groups' attitudes and experiences while participating. An important difference between our study and Azevedo's is that both the self-reported anxiety scores post-task preparation and the custom anxiety question were non-significant in our study, while they were significant in Azevedo et al.'s study. The results were in line with psychophysical data from both studies.

Gathered from the data, the Doppel device did not yield enough of a calming effect to be noticed introspectively by the participants.

5.2 HYPOTHESIS 1

5.2.1 Arousal effects

To investigate the arousal effects attained in the replication of the research done by Azevedo et al. (2017), the main effect of arousal with possible mitigation by the tactile heartbeat emulating device was examined. In our replication, we hypothesized that the application of a heartbeat emulating device would lower overall anxiety when faced with a stressful task.

Across the two conditions, the task preparation was deemed sufficiently stressful, with no significant reduction in skin conductance response to deem the effect significant in the test condition. In our findings, there were no significant effects for any of the physical or psychological measures taken, making our findings at odds with the findings of Azevedo et al.(2017). In the original study, a 20% decrease in heart rate as a setting for the haptic device was enough to elicit a smaller arousal response, a finding we could not replicate.

There are numerous examples found in literature where entrainment between an external stimulus and a somatic system occurs, resulting in anxiety reduction in several different ways, namely in the auditory domain (Ossebaard, 2000);(Casciaro et al., 2013). Entrainment due to a haptic stimulus is severely underused in the literature, which makes a comparison hard to make. While there is evidence that entrainment can occur when people are tasked with a aural stimulus like music or a rhythm(Molinari, Leggio, De Martin, Cerasa, & Thaut, 2003), a haptic stimulus when applied to the wrist does not seem to yield the same results.

It should be noted that both this and the study by Azevedo et al. (2017) had a low amount of participants, one key finding that might explain the difference in significant findings. In our research, we looked at one other possible factor that, in the case of a significant effect, could have explained the difference in findings between this study and the study of Azevedo et al. (2017). Unfortunately, since there is an absence of a main effect, the extra explanations can be considered moot.

5.2.2 Heart Rate Variability and Arousal

With no significant results in both before and after the task, even for the task response, we must conclude that the stress response elicited by the task is not noticeable in the variability of participants' heart rate. What is of note here is the fact that pre- and post-task preparation showed no significant difference in hear rate variability, across the conditions, a finding that is not in line with the remainder of physiological measurements taken.

This could have to do with the fact that the stressful task was relatively short in nature. Some tasks have been unable to elicit a sufficiently large change in physiological responses when they are relatively short in duration (Kristiansen et al., 2009). Moreover, research has shown that the RMSSD is to be sensitive to the High Frequency components of the cardiovascular signal (Berntson, Lozano, & Chen, 2005), and it could be that the lower frequency components show a significant difference in heart rate variability.

5.2.3 Qualitative results

In the study we explored several other avenues which might explain (absence of) the findings by ways of a small qualitative questionnaire. These results revealed the following findings.

The task preparation's goals was obfuscated enough for people to believe they actually needed to give a speech, an aspect that was crucial for the replication to have validity. Almost everyone had no clue as to what the actual intent was of the test and the Doppel device.

The majority of participants did not notice the watch at all, but those who did found it to be annoying. This touches the heart of why the device might not work in its current setting. Participants were required to wear electrodes on their bodies, in particular their non-dominant hand. This hand was also the recipient of the Doppel device and thus remained fairly still and flat on the hard surface of a desk. While the vast majority did not find the electrodes to be disturbing, the Doppel device humming and resonating with the desk was in some cases considered to be especially annoying to the point that people found it hard to concentrate and/or stay calm. In these cases, the Doppel seems to be a stressor instead of a stress-reducing device, influencing the way in which the task at hand was experienced.

These findings make hypothesis #1 "*The usage of a functioning tactile heartbeat simulating device makes people feel less aroused by having a lower skin conductance and higher Heart Rate Variability than when they are not using a functioning tactile heartbeat simulating device"* unproven.

5.3 HYPOTHESIS 2

5.3.1 Entrainment effect

It was hypothesized that the to be observed calming effect would be due to the heart entraining itself to an external haptic stimulus. However, the task preparation provided no significant effect on participants' heartrate. The effect was in the hypothesized direction, meaning that heart rate did slow down for the test group, but not enough to reach significance. These findings are in line with what Azevedo et al. (2017) found, since their work also did not yield an entraining response. Based on this data, we can find no support for hypothesis #2: *Entrainment of the heart and a tactile heartbeat inducing device occurs when people are stimulated by said device.* Lack of entrainment could not be the basis for the calming effect that Azevedo et al. (2017) have found, something that does not follow from our data. Another explanation for the lack of arousal and entrainment could be the degree of interoception that participants might have.

5.4 HYPOTHESIS 3

5.4.1.1 Interoception and arousal

For hypothesis #3: *People that have a high sense of interoception react less strongly to the stimulation of a tactile heartbeat inducing device;* we found a significant effect, showing tentative support for our hypothesis. However, we used a very small (non-normally distributed) subsample, of which no participant formally met the bar for them to be considered "highly interoceptive" as defined by Schandry et al. (2005), leading to a result that is difficult to interpret, but worth exploring in future research. One explanation for participants to not be very interoceptive could be that they were under a tremendous amount of stress a moment before, thus fatiguing them and making it harder to detect bodily processes (Calì, Ambrosini, Picconi, Mehling, & Committeri, 2015). It could also be due to the fact that the heartbeat detection task, while widely used and easy to perform, is also criticized for having low construct validity (Zamariola, Maurage, Luminet, & Corneille, 2018).

5.5 REPLICATION

This study aimed to be a *close replication*, as defined by (Brandt et al., 2014), keeping to the original study for the main effect as close as possible, while adding elements after the main experiment has concluded. Brandt et al. (2014) created a way to determine whether the replication was as close to the original as possible, enabling the researcher to draw firmer conclusions about the validity of the original research and its replication. The method consists of 4 steps or "ingredients" that make up a *replication recipe*:

1. Carefully defining the effects and methods that the researcher intends to replicate;
2. Following exactly the methods of the study;
3. Having high statistical power;
4. Making complete details about the replication available.

In the paper by Brandt et al. (2014), a list of questions was compiled, helping to answer the above questions. Our research aimed to replicate the original study in its entirety, striving to replicate the following effects: *The Doppel device active group will display lower increases in physiological arousal, heartrate entrainment to a haptic frequency and report lower levels of state anxiety.* After performing the instructions by Azevedo et al.(2017) we added extra questions and a heartbeat detection task, which is a deviation from the original study. However, due to the fact that it was added after the formal replication experiment was concluded, this was deemed not to be a factor that might influence the measurements.

Following the exact methods to the letter poses a problem, since there are several aspects that we were unable to replicate, namely:

- Exact location and setting;
- Exact type of participants, mainly in age ranges and gender-mix;
- Exact wording and tone of instructions;

These aspects of the original could not replicated due to physical constraints and a lack of instructions by the study. Extra materials describing these procedures in more detail were not available. Where possible, we followed the wording of the instructions, but no exact playbook existed. Our stimulus was proven to be sufficiently stressful, but do to the lack of a significant effect we wonder whether the placement of the watch and subsequent hand/arm placement was similar to the original study. In our qualitative findings we noticed some participants in the test group were annoyed by the sensation created by the watch resonating with the table, something that may be avoided if the exact placement was mentioned in the methods section.

Azevedo et al. (2017) described the majority of methods for acquiring data and analyzing the results, such as properly defined questionnaires and pre-processing methods. The pre-processing process of the SC-AUC signal transformation was not documented however, leaving us to select a filter that may or may not have been used in the original study. Bach et al., (2010) described the pre-processing process, but leaves open the selection of filters, further complicating the replication of Azevedo et al. (2017) exact signal pre-processing. Though a similar effect has been acquired in the averaging of the signal, a properly replicated signal-transformation method could be contributed to a more similar signal and thus results.

One last aspect that makes our study differ from the original, pertains to the physical recording device that acquires data. Though there are certainly industry standards that make the acquiring of biophysical signals of similar quality, our device differed from the one used by Azevedo et al.(2017).

Having a statistical power large enough to detect effects is key in the replication process. Brandt et al. (2014) quotes a number of studies that list power between 0.8 and 0.95 and/or using a sample size that is roughly 2.5 times larger than the study one tries to replicate. We have opted for a power of 0.9, in line with our department practice, along with a sequential analysis, putting us in the middle of the guidelines. The power of the original study is not listed, but could have been 0.9 – 0.99, based on the effect size and sample size reported, making our study sufficiently similar.

All gathered data can be found at the online data repository of Eindhoven University of Technology, along with a copy of this report.

5.6 EXPERIMENT CONTINUATION

Prior to the experiment, we defined two metrics which were the basis for continuing the research to reach a Cohen's D of 0.54:

A) The likely effect size exceeds 0.3;
B) The P-value of the t-test on the skin conductance response > 0.031;

The initial power for the complete experiment with two groups of 60 participants was 0.9, which is reduced to 0.7 for two groups of 30 with an effect size of 0.53 as per Azevedo et al.(2017). Conversely, our study is able to detect an effect size of Cohen's D = 0.76 with a power of 0.9, based on the largest effect size gathered from the original study via independent t-tests. The same analysis yielded a Cohen's D = 0.44 for the AUC Skin Conductance response, with a p = 0.095(Lakens, 2013).

Our findings show that criterion B is attained due to the results being non-significant and being pretty far from being significant. Criterion A is likely met based on Cohen's D = 0.442 for the t-test. It is likely that with this effect size, the original projection for the amount of participants is not sufficient, making termination of the current research necessary based on criteria stated in paragraph 3.6.6.

It should be noted that for this initial power analysis, we took the largest effect size that Azevedo et al.(2017) reported for the independent t-tests as an after analysis. When the partial eta squared for the SC-AUC-interaction effect TIME*CONDITON in the original study is taken, a different picture emerges. Azevedo et al.(2017) reported a partial $\eta^2 = 0.095$. Follow-up sensitivity analysis shows that with an alpha level of 0.05, a sample size of 0.9 and an effect size of 0.32 (following from the partial $\eta^2 = 0.095$) we would have a power of 0.995, implying that the effect, if present, should be detectable. If we apply the same analysis to our findings for the SC-AUC data, alpha = 0.05, partial $\eta^2 = 0.022$, participants = 60, we get a power of 0.627, based on an effect size f of 0.147. An a-priori analysis shows that to reliably detect an effect of that magnitude, you would need the original 120 participants. Criterion B is, when comparing the magnitude of the *f* – score for Cohen's *d* (which was chosen to be a medium effect*)* (Lakens, 2013) unlikely to be met or exceeded based on our data, and we would recommend suspending further analysis due to the very small effect elicited by the Doppel device.

5.7 LIMITATIONS AND FUTURE RESEARCH

The achieved power of this study is 0.67 for the AUC Skin Conductance Response with its current amount of participants, making detection of an effect less likely. This could be one of the reasons that there was no significant effect found for the Doppel study.

Limitations described in our replication paragraph apply, calling for a better documentation of the former and possible future studies to better compare individual methods and results.

The absence of a main effect could have been due to people finding the way in which the Doppel was mounted on their wrist uncomfortable. In our findings, we have seen several people complaining about noticing the wrist mounted device while resting their wrist on a table. It could be that when people do the experiment while standing or holding their wrist in another manner, that the Doppel becomes less of a stressor and more of a calming device.

The question why entrainment does not take place or why people do not get calmed in general by applying a calming device to the skin is one that warrants more research. The different pathways in which people are calmed by facing a rhythmic heartbeat emulating device is an avenue worth pursuing. This particular paper focused on checking for an actual effect, with limited supplemental data for explaining (the absence of) an effect. One area which is unexplored as of yet is comparing two stimuli for their relative effect. One could think of an aural and haptic stimulus, similar in frequency and sensation, checking whether people are calmed by heartbeat emulating devices in general. The effect of music is widely recognized, but does this effect translate to more primal aural stimuli?

Another area that could yield interesting results is seeing which neurophysiological mechanisms underpin the calming reaction. When a calming effect can be established using a haptic device, the activated neural pathways responsible for influencing our homeostasis and related stress response could be further examined, laying bare whether or not this is due to actual cardiac entrainment or the brain mistaking an eternal stimulus for the heart.

Our study had too few participants to say with certainty whether a small effect could be there or not. Future studies, with the above recommendations, could yield interesting results when coupled with more participants.

5.8 CONCLUSION

In this paper, we set out to discover whether the usage of a tactile heartbeat emulating stimulus would result in a lower stress response due to cardiac entrainment. We did this by ways of applying a watch-like device to participants' wrists, vibrating at a frequency 20% lower than their resting heartrate, while presenting them with a stressful task.

The results were that the task was significantly stressful enough. However, we found no beneficial arousal reduction effect of the heartbeat-like stimulations across all taken measurements, be it self-reported subjective anxiety, Skin Conductance Averges, Skin Conductance-AUC and entrainment.

The to be perceived effect is likely to be relatively small, making detection possible with a larger sample size. However, having observed the likely effect size to be small, we would advise against continuing the same study with the added participants. In this paper we provided a number of recommendations to continue research into calming effect of haptic heartbeat inducing stimuli since they are a relatively simple way of calming people. The societal benefits could be significant if the product was tailored towards all circumstances in which people wear watches, since this makes external validity of the product higher. Exploring different contexts in which these kinds of devices could be used opens up more avenues of research, making detection of an effect more likely, aiding people in keeping calm across different contexts.

REFERENCES

Aarts, A. A., Anderson, J. E., Anderson, C. J., Attridge, P. R., Attwood, A., Axt, J., … Zuni, K. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). https://doi.org/10.1126/science.aac4716

Allen, J. J. B. (2002). Calculating metrics of cardiac chronotropy: A pragmatic overview. *Psychophysiology*.

Azevedo, R. T., Bennett, N., Bilicki, A., Hooper, J., Markopoulou, F., & Tsakiris, M. (2017). The calming effect of a new wearable device during the anticipation of public speech. *Scientific Reports*, *7*(1), 1–7. https://doi.org/10.1038/s41598-017-02274-2

Bach, D. R., Friston, K. J., & Dolan, R. J. (2010). Analytic measures for quantification of arousal from spontaneous skin conductance fluctuations. *International Journal of Psychophysiology*, *76*(1), 52–55. https://doi.org/10.1016/j.ijpsycho.2010.01.011

Bernardi, L., Porta, C., & Sleight, P. (2006). Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: The importance of silence. *Heart*. https://doi.org/10.1136/hrt.2005.064600

Berntson, G. G., Lozano, D. L., & Chen, Y. J. (2005). Filter properties of root mean square successive difference (RMSSD) for heart rate. *Psychophysiology*. https://doi.org/10.1111/j.1469-8986.2005.00277.x

Boschman, M. C. (2012). Acquisition of Physiological Data @ HTI - User Manual, (November).

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., … van ’t Veer, A. (2014). The Replication Recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, *50*(1), 217–224. https://doi.org/10.1016/j.jesp.2013.10.005

Burman, L. E., Reed, W. R., & Alm, J. (2010). A call for replication studies. *Public Finance Review*. https://doi.org/10.1177/1091142110385210

Calì, G., Ambrosini, E., Picconi, L., Mehling, W. E., & Committeri, G. (2015). Investigating the relationship between interoceptive accuracy,interoceptive awareness,and emotional susceptibility. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2015.01202

Canolty, R. T., & Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2010.09.001

Carleton, R. N., McCreary, D. R., Norton, P. J., & Asmundson, G. J. G. (2006). Brief fear of negative evaluation scale - Revised. *Depression and Anxiety*. https://doi.org/10.1002/da.20142

Casciaro, F., Laterza, V., Conte, S., Pieralice, M., Federici, A., Todarello, O., … Conte, E. (2013). Alpha-rhythm stimulation using brain entrainment enhances heart rate variability in subjects with reduced HRV. *World Journal of*

Neuroscience. https://doi.org/10.4236/wjns.2013.34028

Chrousos, G. P. (2009, July). Stress and disorders of the stress system. *Nature Reviews Endocrinology*. https://doi.org/10.1038/nrendo.2009.106

Chrousos, G. P., & Gold, P. W. (1992). The Concepts of Stress and Stress System Disorders: Overview of Physical and Behavioral Homeostasis. *JAMA: The Journal of the American Medical Association*. https://doi.org/10.1001/jama.1992.03480090092034

Clayton, M., Sager, R., & Will, U. (2005). In time with the music : the concept of entrainment and its significance for ethnomusicology. *European Meetings in Ethnomusicology.*

Clemens, W. J. (1984). Temporal Arrangement of Signals in Heartbeat Discrimination Procedures. *Psychophysiology*. https://doi.org/10.1111/j.1469-8986.1984.tb00202.x

Craig, A. D. (2003). Interoception: The sense of the physiological condition of the body. *Current Opinion in Neurobiology*. https://doi.org/10.1016/S0959-4388(03)00090-4

Dale, A., & Anderson, D. (1978). Information variables in voluntary control and classical conditioning of heart rate: Field dependence and heart-rate perception. *Perceptual and Motor Skills*. https://doi.org/10.2466/pms.1978.47.1.79

Dawson, M. E., Schell, A. M., & Filion, D. L. (2016). The electrodermal system. In *Handbook of Psychophysiology, Fourth Edition*. https://doi.org/10.1017/9781107415782.010

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*. https://doi.org/10.1037/0033-2909.130.3.355

Dijk, E. O., & Weffers, A. (2013). Systems to support relaxation by using tactile, acoustic and visual stimuli. *International Journal of Autonomous and Adaptive Communications Systems*, *6*(4), 324. https://doi.org/10.1504/IJAACS.2013.056819

Dilworth, C. (1994). Principles, laws, theories and the metaphysics of science. *Synthese*. https://doi.org/10.1007/BF01064018

Etzel, J. A., Johnsen, E. L., Dickerson, J., Tranel, D., & Adolphs, R. (2006). Cardiovascular and respiratory responses during musical mood induction. *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2005.10.025

Everly, G. S., & Lating, J. M. (2012). *A Clinical Guide to the Treatment of the Human Stress Response*. *A Clinical Guide to the Treatment of the Human Stress Response*. https://doi.org/10.1007/978-1-4614-5538-7

Fredrikson, M., & Matthews, K. A. (1990). Cardiovascular responses to behavioral stress and hypertension: A meta-analytic review. *Annals of Behavioral Medicine*. https://doi.org/10.1207/s15324796abm1201_3

Freese, J., & Peterson, D. (2017). *Replication in Social Science*. *SSRN*. https://doi.org/10.1146/annurev-soc-060116-053450

Gallace, A., & Spence, C. (2010). The science of interpersonal touch: An overview. *Neuroscience and Biobehavioral Reviews*, *34*(2), 246–259. https://doi.org/10.1016/j.neubiorev.2008.10.004

Garfinkel, S. N., Seth, A. K., Barrett, A. B., Suzuki, K., & Critchley, H. D. (2015). Knowing your own heart: Distinguishing interoceptive accuracy from interoceptive awareness. *Biological Psychology*, *104*, 65–74. https://doi.org/10.1016/j.biopsycho.2014.11.004

Golombek, D. A., & Rosenstein, R. E. (2010). Physiology of circadian entrainment. *Physiological Reviews*. https://doi.org/10.1152/physrev.00009.2009

Haas, F., Distenfeld, S., & Axen, K. (1986). Effects of perceived musical rhythm on respiratory pattern. *Journal of Applied Physiology*. https://doi.org/10.1152/jappl.1986.61.3.1185

Harms, C., & Lakens, D. (2018). Making “null effects” informative: statistical techniques and inferential frameworks. *Journal of Clinical and Translational Research*, *3*, 382–393. https://doi.org/10.18053/jctres.03.2017s2.007

Hellhammer, J., & Schubert, M. (2012). The physiological response to Trier Social Stress Test relates to subjective measures of stress during but not before or after the test. *Psychoneuroendocrinology*. https://doi.org/10.1016/j.psyneuen.2011.05.012

Hempel, C. G. (1968). Maximal Specificity and Lawlikeness in Probabilistic Explanation. *Philosophy of Science*. https://doi.org/10.1086/288197

Joëls, M., & Baram, T. Z. (2009). The neuro-symphony of stress. *Nature Reviews Neuroscience*. https://doi.org/10.1038/nrn2632

Julian, L. J. (2011). Measures of anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A). *Arthritis Care and Research*, *63*(SUPPL. 11), 467–472. https://doi.org/10.1002/acr.20561

Khalfa, S., Roy, M., Rainville, P., Dalla Bella, S., & Peretz, I. (2008). Role of tempo entrainment in psychophysiological differentiation of happy and sad music? *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2007.12.001

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The Evolution of Rhythm Processing. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2018.08.002

Kristiansen, J., Mathiesen, L., Nielsen, P. K., Hansen, Å. M., Shibuya, H., Petersen, H. M., … Søgaard, K. (2009). Stress reactions to cognitively demanding tasks and open-plan office noise. *International Archives of Occupational and Environmental Health*, *82*(5), 631–641. https://doi.org/10.1007/s00420-008-0367-4

Kudielka, B. M., Wüst, S., Kirschbaum, C., & Hellhammer, D. H. (2007). Trier Social Stress Test. In *Encyclopedia of Stress*. https://doi.org/10.1016/B978-012373947-6.00681-4

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*. https://doi.org/10.1126/science.1154735

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2013.00863

Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses. *European Journal of Social Psychology*. https://doi.org/10.1002/ejsp.2023

Leary, M. R. (1983). A Brief Version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin*, *9*(3), 371–375. https://doi.org/10.1177/0146167283093007

Lidberg, L., & Wallin, B. G. (1981). Sympathetic Skin Nerve Discharges in Relation to Amplitude of Skin Resistance Responses. *Psychophysiology*. https://doi.org/10.1111/j.1469-8986.1981.tb03033.x

Mai, J. K., & Paxinos, G. (2012). *The Human Nervous System*. *The Human Nervous System*. https://doi.org/10.1016/C2009-0-02721-4

McCarty, R. (2016). The Fight-or-Flight Response: A Cornerstone of Stress Research. In *Stress: Concepts, Cognition, Emotion, and Behavior: Handbook of Stress*. https://doi.org/10.1016/B978-0-12-800951-2.00004-2

McGuigan, F. J., & Andreassi, J. L. (1981). Psychophysiology -- Human Behavior and Physiological Response. *The American Journal of Psychology*. https://doi.org/10.2307/1422751

McPherson, C. K., & Armitage, P. (1971). Repeated Significance Tests on Accumulating Data when the Null Hypothesis is not True. *J. R. Stat. Soc., Ser. A*. https://doi.org/10.1135/cccc19930629

Melzig, C. A., Weike, A. I., Hamm, A. O., & Thayer, J. F. (2009). Individual differences in fear-potentiated startle as a function of resting heart rate variability: Implications for panic disorder. *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2008.07.013

Menon, G. K., & Kligman, A. M. (2009). Barrier functions of human skin: A holistic view. *Skin Pharmacology and Physiology*. https://doi.org/10.1159/000231523

Miller, S. M. (1979). Controllability and human stress: Method, evidence and theory. *Behaviour Research and Therapy*, *17*(4), 287–304. https://doi.org/10.1016/0005-7967(79)90001-9

Molinari, M., Leggio, M. G., De Martin, M., Cerasa, A., & Thaut, M. (2003). Neurobiology of Rhythmic Motor Entrainment. *Annals of the New York Academy of Sciences*, *999*, 313–321.

https://doi.org/10.1196/annals.1284.042

Murayama, K., Pekrun, R., & Fiedler, K. (2014). Research Practices That Can Prevent an Inflation of False-Positive Rates. *Personality and Social Psychology Review*, *18*(2), 107–118. https://doi.org/10.1177/1088868313496330

Mütze, H., Kopiez, R., & Wolf, A. (2018). The effect of a rhythmic pulse on the heart rate: Little evidence for rhythmical 'entrainment' and 'synchronization.' *Musicae Scientiae*. https://doi.org/10.1177/1029864918817805

Ossebaard, H. C. (2000). Stress reduction by technology? An experimental study into the effects of brainmachines on burnout and state anxiety. *Applied Psychophysiology Biofeedback*. https://doi.org/10.1023/A:1009514824951

Park, S., & Jayaraman, S. (2003). Enhancing the Quality of Life Through Wearable Technology. *IEEE Engineering in Medicine and Biology Magazine*, *22*(3), 41–48. https://doi.org/10.1109/MEMB.2003.1213625

Pereira, T., Almeida, P. R., Cunha, J. P. S., & Aguiar, A. (2017). Heart rate variability metrics for fine-grained stress level assessment. *Computer Methods and Programs in Biomedicine*. https://doi.org/10.1016/j.cmpb.2017.06.018

Phillips-Silver, J., Aktipis, C. A., & A. Bryant, G. (2010). The Ecology of Entrainment: Foundations of Coordinated Rhythmic Movement. *Music Perception*, *28*(1), 3–14. https://doi.org/10.1525/mp.2010.28.1.3

Pocock, S. J. (1977). Group sequential methods in the design and analysis of clinical trials. *Biometrika*. https://doi.org/10.1093/biomet/64.2.191

Pollatos, O., Kirsch, W., & Schandry, R. (2005). On the relationship between interoceptive awareness, emotional experience, and brain processes. *Cognitive Brain Research*, *25*(3), 948–962. https://doi.org/10.1016/j.cogbrainres.2005.09.019

Pollatos, O., & Schandry, R. (2004). Accuracy of heartbeat perception is reflected in the amplitude of the heartbeat-evoked brain potential. *Psychophysiology*, *41*(3), 476–482. https://doi.org/10.1111/1469-8986.2004.00170.x

Prinsloo, G. E., Derman, W. E., Lambert, M. I., & Laurie Rauch, H. G. (2013). The effect of a single episode of short duration heart rate variability biofeedback on measures of anxiety and relaxation states. *International Journal of Stress Management*. https://doi.org/10.1037/a0034777

Rashidi, P., & Mihailidis, A. (2013). A survey on ambient-assisted living tools for older adults. *IEEE Journal of Biomedical and Health Informatics*. https://doi.org/10.1109/JBHI.2012.2234129

Sanderson, P. M., & Fisher, C. (1994). Exploratory Sequential Data Analysis: Foundations. *Humanâ��Computer Interaction*. https://doi.org/10.1080/07370024.1994.9667208

Saul, J. (1990). Beat-To-Beat Variations of Heart Rate Reflect Modulation of Cardiac Autonomic Outflow. *Physiology*.

https://doi.org/10.1152/physiologyonline.1990.5.1.32

Schandry, R. (1981). Heart Beat Perception and Emotional Experience. *Psychophysiology*, *18*(4), 483–488. https://doi.org/10.1111/j.1469-8986.1981.tb02486.x

Schmidt, S. (2009). Shall We Really Do It Again? The Powerful Concept of Replication Is Neglected in the Social Sciences. *Review of General Psychology*, *13*(2), 90–100. https://doi.org/10.1037/a0015108

Seidel, H., & Herzel, H. (1998). Analyzing entrainment of heartbeat and respiration with surrogates. *IEEE Engineering in Medicine and Biology Magazine*, *17*(6), 54–57. https://doi.org/10.1109/51.731321

Serdaroglu, K., Uslu, G., & Baydere, S. (2015). Medication intake adherence with real time activity recognition on IoT. *2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications, WiMob 2015*, 230–237. https://doi.org/10.1109/WiMOB.2015.7347966

Shaffer, F., McCraty, R., & Zerr, C. L. (2014). A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. *Frontiers in Psychology*, *5*(September), 1–19. https://doi.org/10.3389/fpsyg.2014.01040

Smith, C. A. (1989). Dimensions of appraisal and physiological response in emotion. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.56.3.339

Spielberger, C. D., Gorsuch, R. L., Lushene, P. R., Vagg, P. R., & Jacobs, A. G. (1983). *Manual for the State-Trait Anxiety Inventory (Form Y)*. *Manual for the statetrait anxiety inventory STAI*. https://doi.org/10.1007/978-1-4419-9893-4

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (2010). State-Trait Anxiety Inventory for Adults Sampler Set Manual, Instrument and Scoring Guide, 0–75. Retrieved from www.mindgarden.com.

Streiner, D. L. (1993). An introduction to multivariate statistics. *Canadian Journal of Psychiatry*. https://doi.org/10.1177/070674379303800104

Swan, M. (2012). Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0. *Journal of Sensor and Actuator Networks*. https://doi.org/10.3390/jsan1030217

Thaut, M. H., Tian, B., & Azimi-Sadjadi, M. R. (1998). Rhythmic finger tapping to cosine-wave modulated metronome sequences: Evidence of subliminal entrainment. *Human Movement Science*. https://doi.org/10.1016/S0167-9457(98)00031-1

Thayer, J. F., Åhs, F., Fredrikson, M., Sollers, J. J., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health. *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2011.11.009

Tierney, A., & Kraus, N. (2014). Neural entrainment to the rhythmic structure of music. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn_a_00704

Watanabe, T., Okubo, M., & Kuroda, T. (1996). Analysis of entrainment in face-to-face interaction using heart rate variability. In *Robot and Human Communication - Proceedings of the IEEE International Workshop*. https://doi.org/10.1109/roman.1996.568786

Werner, N. S., Duschek, S., Mattern, M., & Schandry, R. (2009). Interoceptive sensitivity modulates anxiety during public speaking. *Journal of Psychophysiology*. https://doi.org/10.1027/0269-8803.23.2.85

Westenberg, P. M., Bokhorst, C. L., Miers, A. C., Sumter, S. R., Kallen, V. L., van Pelt, J., & Blöte, A. W. (2009). A prepared speech in front of a pre-recorded audience: Subjective, physiological, and neuroendocrine responses to the Leiden Public Speaking Task. *Biological Psychology*. https://doi.org/10.1016/j.biopsycho.2009.06.005

Whitehead, W. E., Drescher, V. M., Heiman, P., & Blackwell, B. (1977). Relation of heart rate control to heartbeat perception. *Biofeedback and Self-Regulation*, *2*(4), 371–392. https://doi.org/10.1007/BF00998623

Wiens, S. (2005). Interoception in emotional experience. *Current Opinion in Neurology*. https://doi.org/10.1097/01.wco.0000168079.92106.99

Wilson, M., & Cook, P. F. (2016). Rhythmic entrainment: Why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed. *Psychonomic Bulletin and Review*. https://doi.org/10.3758/s13423-016-1013-x

Zamariola, G., Maurage, P., Luminet, O., & Corneille, O. (2018). Interoceptive accuracy scores from the heartbeat counting task are problematic: Evidence from simple bivariate correlations. *Biological Psychology*. https://doi.org/10.1016/j.biopsycho.2018.06.006

6 APPENDIX A: QUANTITATIVE & QUALITATIVE QUESTIONNAIRES

6.1 STAI-Y1

Participant #…………………………………………… Date……….… c / t pre

DIRECTIONS: Several statements which people have used to describe themselves are given below. Read each statement and then circle the number that indicates how you feel right now, that is, at this moment. There is no right or wrong answer. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

#		Not at all	Somewhat	Moderately So	Very much so
1.	I feel calm	1	2	3	4
2.	I feel secure	1	2	3	4
3.	I am tense	1	2	3	4
4.	I feel strained	1	2	3	4
5.	I feel at ease	1	2	3	4
6.	I feel upset	1	2	3	4
7.	I am presently worrying over possible misfortunes	1	2	3	4
8.	I feel satisfied	1	2	3	4
9.	I feel frightened	1	2	3	4
10.	I feel comfortable	1	2	3	4
11	I feel self confident	1	2	3	4
12.	I feel nervous	1	2	3	4
13.	I am jittery	1	2	3	4
14.	I feel indecisive	1	2	3	4
15.	I am relaxed	1	2	3	4
16.	I feel content	1	2	3	4
17.	I am worried	1	2	3	4
18.	I feel confused	1	2	3	4
19.	I feel steady	1	2	3	4
20.	I feel pleasant	1	2	3	4

6.2 BFNE

Participant ID #:_________________

Date:__________________

6.2.1.1.1 *Please circle the number that best corresponds to how much you agree with each item.*

	Not at all characteristic of me	A little characteristic of me	Somewhat characteristic of me	Very characteristic of me	Entirely characteristic of me
1. I worry about what other people will think of me even when I know it doesn't make any difference.	1	2	3	4	5
2. It bothers me when people form an unfavourable impression of me.	1	2	3	4	5
3. I am frequently afraid of other people noticing my shortcomings.	1	2	3	4	5
4. I worry about what kind of impression I make on people.	1	2	3	4	5
5. I am afraid that others will not approve of me.	1	2	3	4	5
6. I am afraid that other people will find fault with me.	1	2	3	4	5
7. I am concerned about other people's opinions of me.	1	2	3	4	5
8. When I am talking to someone, I worry about what they may be thinking about me.	1	2	3	4	5
9. I am usually worried about what kind of impression I make.	1	2	3	4	5
10. If I know someone is judging me, it tends to bother me.	1	2	3	4	5
11. Sometimes I think I am too concerned with what other people think of me.	1	2	3	4	5
12. I often worry that I will say or do wrong things.	1	2	3	4	5

6.3 AZEVEDO QUESTIONNAIRE

Participant #:

Please indicate by circling the number how you feel about the following statements:

	Strongly disagree							Strongly agree
I like to speak in public	1	2	3	4	5	6	7	8
I can easily become anxious in social situations	1	2	3	4	5	6	7	8

	Stressful						Relaxing
My experience during the task preparation was:	1	2	3	4	5	6	7
	Unpleasant						Pleasant
My experience during the task preparation was:	1	2	3	4	5	6	7

6.5 DEDICATED QUESTIONNAIRE AFTER EXPERIMENT

Participant #: Date:

Please fill in this questionnaire after completing the previous 3:

- STAY-Y1;
- bFNE;
- Azevedo Questionnaire;

Below are some open and closed questions, please fill these in to the best of your ability.

#	Question
1	How did you feel about giving a speech to an unknown audience?
Answer:	
2	Did you believe you actually needed to give the speech?
Answer:	
3	Were you aware of the watch during the test?
Answer:	
4	Did it influence the way in which you conducted the tests?
Answer:	
5	Did you try to keep calm in a deliberate manner?

Answer: **SEE OTHER SIDE!**	
6	Did you feel calmed by the vibrating nature of the device?
Answer:	
7	You were hooked up to different devices. Did the devices feel comfortable?
Answer:	
8	Can you imagine getting used to wearing the watch regularly?
Answer:	
9	Can you imagine getting used to wearing the watch when you wear it an entire day?
Answer:	
10	Were you aware of the actual intent of the watch during the test?
Answer:	
11	Did it influence the way in which you conducted the tests?
Answer:	

6.6 SCHANDRY HEARTBEAT DETECTION

Please fill in the amount of heartbeats you detected in the intervals indicated by the researcher.

Interval Number	Amount of heartbeats detected	
1		
2		
3		

Please fill in the time in seconds you think has passed in the intervals indicated by the researcher.

Interval Number	Time passed in seconds
1	
2	
3	

7 APPENDIX B: INFORMED CONSENT FORM

This document gives you information about the study "Blood Pressure Measure". Before the study begins, it is important that you learn about the procedure followed in this study and that you give your informed consent for voluntary participation. Please read this document carefully.

Aim and benefit of the study

The aim of this research is to measure the quality of a new blood pressure monitor. The information obtained is used to evaluate and possibly improve the design of this meter.

This study is performed by Geert van der Velden, a student under the supervision of Prof. Dr. J. Westerink of the Human-Technology Interaction group.

Procedure

In this experiment you will be seated in a chair with 3 measuring instruments, namely:
- 3 electrodes directly on your body, of which 2 on the chest;
- A measuring instrument on the middle finger of the hand that you do not write with;
- a blood pressure measuring watch;

After these instruments have been tested you will do a short test of 5 minutes. You will complete a short survey prior to this test. A rundown of the test follows after the application and testing of the above instruments is completed.

After this test has ended, you fill in 3 more short surveys. The final part is a short test of about 1 minute. You can remain seated; all the necessary material is within reach.

After completing the above, the executor will disconnect the devise from your body. Then you will receive a debriefing shortly afterwards.

Risks

The study does not involve any risks or detrimental side effects.

Duration

The study will last approximately 45-60 minutes.

Participants

You were selected because you were registered as participant in the participant database of the Human Technology Interaction group of the Eindhoven University of Technology. It might also be the case that you are selected by a personal invitation from the experimenter.

Voluntary

Your participation is completely voluntary. You can refuse to participate without giving any reasons and you can stop your participation at any time during the study. You can also withdraw your permission to use your data up to 24 hours after the study is finished. All this will have no negative consequences whatsoever.

Compensation

You will be paid 10 euros (€2.00 extra if you do not study or work at the TU/e or Fontys Eindhoven).

Confidentiality

All research conducted at the Human-Technology Interaction Group adheres to the Code of Ethics of the NIP (Nederlands Instituut voor Psychologen – Dutch Institute for Psychologists).

We will not be sharing personal information about you to anyone outside of the research team. No video or audio recordings are made that could identify you. The information that we collect from this study is used for writing scientific publications and will only be reported at group level. It will be completely anonymous and it cannot be traced back to you.

Further information

If you want more information about this study you can ask Geert van der Velden, g.p.w.v.d.velden@student.tue.nl.

If you have any complaints about this study, please contact the supervisor, Prof. Dr. J. Westerink, j.h.d.m.westerink@tue.nl.

Certificate of Consent

I, (NAME)……………………………………….. have read and understood this consent form and have been given the opportunity to ask questions. I agree to voluntarily participate in this study carried out by the research group Human Technology Interaction of the Eindhoven University of Technology.

Participant's Signature	Date