UNIVERSITY
OF TWENTE.

Final Version - Master Thesis

Exploring the relationship between Residuals from Acceleration predicted Heart Rate (RAHR) and self-reported momentary and retrospective stress and arousal

Jule Krüger, s1408798

Master Human Factors and Engineering Psychology

16-01-2018

Supervisors:

1st: Dr. Matthijs Noordzij, Department of Psychology, Health & Technology

2nd: Erika van Lier, Department of Cognitive Psychology & Ergonomics

University of Twente

Faculty of Behavioural, Management, and Social Sciences

Department of Cognitive Psychology and Ergonomics

Abstract

As more and more people are buying wearables and are interested in monitoring their stress level with them, it gets increasingly important to find a good way of inferring psychological measures such as stress and arousal from physiologically measurable variables such as heart rate. As laboratory experiments may not be able to predict relationships in real life, a wearable and experience sampling study taking place in daily life was chosen for the current study. Because heart rate changes are associated with both a physical and a psychological part, it is important to try to account for the physical part when the interest lies in the heart rate changes mostly associated with psychological changes. Previously, there have been different approaches to do this, which are all not practical for a study in daily life. In the present study, a new approach to calculating the psychologically caused part of heart rate called “Residuals from Acceleration predicted Heart Rate” (RAHR) is explored. In this approach, a regression is made for heart rate predicted by magnitude of acceleration (physical activity). The positive residuals of this are taken as RAHR. This approach can be implemented with only a heart rate sensor and an accelerometer in the wearable. A correlation for acceleration and heart rate was found and RAHRs were implemented. Unfortunately, no relationship could be found between RAHR and the four measures of psychological experience, namely momentary and retrospective experienced stress, and momentary and retrospective experienced arousal. Still, a new method for calculating the psychological part of heart rate has been tested which can be used in the field instead of the laboratory. This method now must be validated further in order to see if it may be usable for giving feedback about people’s emotional state. Further research is needed for this.

ACKNOWLEDGEMENT

I would like to thank my thesis supervisors Dr. Matthijs Noordzij and Erika van Lier for giving very useful feedback through the process of my thesis. Their feedback was very helpful and made my text a lot better. The meetings with Matthijs were always refreshing and provided me with answers for my questions and new ideas for my next steps. I furthermore thank him for providing me with the topic for my thesis and the data that had already been collected by other students.

I would really like to thank the students who worked on this topic before me for the seven-days-long data collection. This study would not have been possible without them. Of course, I would also like to thank the participants of this study for taking this much work upon them during data collection.

Finally, I would like to thank my loved ones for providing me with continuous support and encouragement throughout my years of study and throughout the process of writing this thesis. This accomplishment would not have been possible without them. Thank you.

Author
Jule Krüger

Table of Contents

Introduction

The demand for wearables, such as fitness bands, smart bands, or smart watches, is increasing. The prediction is that the market for all wearables will nearly double from 2017 to 2021, with watches and wristbands still holding nearly 90% market share of all types of wearables in 2021 (IDC, 2017). One main goal of an increasing number of those devices is giving users feedback about their stress levels (Maslakovic, 2017). Receiving this feedback is also one of the demands users have for those devices, which can be seen in a survey from the UK (Valencell, 2016). Here, 55% of the participants who owned a wearable would like to monitor their stress with it. Receiving information about your stress level is seen as an impulse for actively trying to relax during high stress phases and thereby making your life healthier and happier. There are already devices that offer solutions for relaxing in addition to this stress-level information (Caddy, 2017).

Stress and other psychological experience (e.g., arousal or a specific emotion such as anger or fear) is often associated with activity in the autonomic nervous system (ANS; Kreibig, 2010). As ANS activity can be measured through different physiological measures, the psychological variable stress can be deduced from different sorts of data. Sandulescu, Andrews, Ellis, Bellotto, and Mozos (2015) for example used heart rate (HR) and electrodermal activity (EDA) to measure the stress level of their participants, while in a study by Wijsman, Grundlehner, Liu, Hermens, and Penders (2011) electrocardiography (ECG), respiration, skin conductance, and electromyography (EMG) were used. Although stress-level recognition is already implemented in some commercial wearables, inferring psychological experience as stress, a specific emotion, or general arousal from physiological measures was found to be quite difficult, as the relation between physiological events (i.e., responses of the body) and psychological events (i.e., experience of the mind) is more complex than lay-people and scientists would initially hypothesize (Evers et al., 2014; Fairclough, 2009; Feldman Barrett, 2006).

Cacioppo, Tassinary, and Berntson (2017) define psychological events as ‘conceptual variables representing functional aspects of embodied processes’ (Cacioppo et al., 2017, p. 8) and physiological events as ‘empirical physical variables’ (Cacioppo et al., 2017, p. 8). They believe that such events should be seen as representing two domains or sets made up of different elements. The elements in one set can be related with the elements in the other set in five different ways (Figure 1). Firstly, in a one-to-one relation, one element from the psychological set is associated with one element from the physiological set, while, secondly, in a many-to-many relation, two or more psychological and physiological elements are associated with each other. Thirdly, in a many-to-one relation, two or more psychological elements are associated with one physiological element, and fourthly, in a one-to-many relation, the opposite is the case. In the fifth sort of relation, a null relation, there is no association between a psychological element and the physiological element (Cacioppo et al., 2017). Although one-to-one relations are the easiest to interpret, they should not be assumed without reason. Evers et al. (2014), presented a dual-process framework for coherence between the experiential and physiological systems and showed that this coherence only takes place within one system, but not across systems. Another assumption that should not be made is that relations between psychological and physiological events are the same across situations (and time) and individuals (Cacioppo et al., 2017). There are differences in experiencing emotions, both within and between situations and individuals (Russell, 2009).

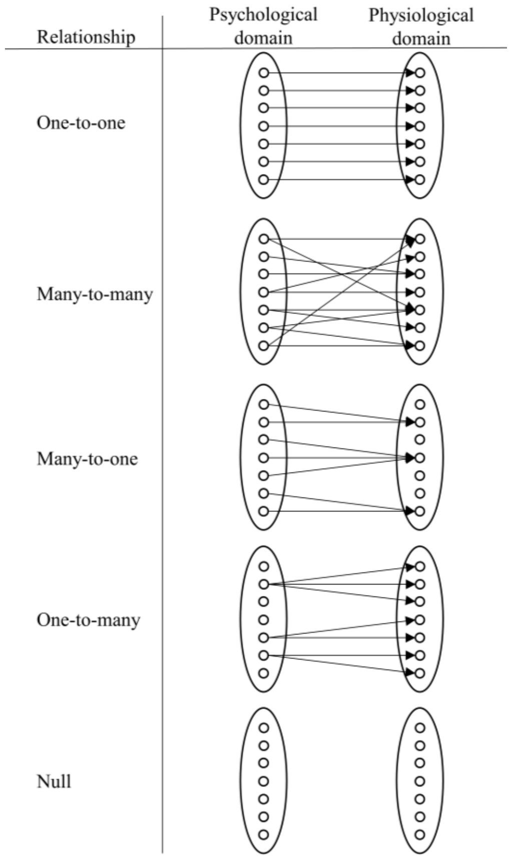

Figure 1. Possible relationships between domains

As the inference of psychological activity from physiological activity is so complex, more research is needed in this domain. While laboratory studies researching stress and emotions may be easier to control, they may not represent real-life situations, both due to the kind of situations used and their predictability in the laboratory versus unpredictability in the field (i.e., real life). Laboratory studies thus lack good ecological validity (Zanstra & Johnston, 2011). Through technological advances, real-life field studies are increasingly easy to execute. Physiological activity (e.g., heart rate, EDA) can be measured more easily and less obtrusively through small, wearable sensors. Also, psychological variables (e.g., stress, emotion, arousal) as measured through self-report can now be retrieved in real time through the use of electronic diary methods, for example by the use of smartphone apps (Zanstra & Johnston, 2011). Using these technological advances makes it easier to correlate psychological and physiological activity. This can be done during real life and in real-time, which increases the ecological validity of the study.

In the current study, the mentioned technological advances are used to explore the relationship between physiological and psychological events in daily life. As ANS activity, in the current study measured through heart rate, is generally said to be influenced by both physical and psychological activity (Carroll, Turner, & Rogers, 1987), a new way to filter out the physical part, called "Residuals from Acceleration predicted Heart Rate" (RAHR), is developed and described below. This way, it is possible to examine which part of the participants' sensor-measured heart rate is not caused physically but psychologically. This part of the heart rate is then correlated to their subjective stress and arousal.

Physiological activity

Because of the availability of wearable sensors, physiological activity can be measured easier, faster, less obtrusively, and continuously in daily life. In the current study, heart rate was chosen as a measure for ANS activity. Measuring heart rate is a feature of many wearable sensors, so that it is readily available for many wearable owners.

Heart Rate. Heart rate is generally said to incorporate both physical activity and psychological activity (Carroll et al., 1987). The physically induced heart rate is evoked by metabolic activity, i.e. oxygen needs to be transported to the relevant muscles faster, leading to an increase in blood-pumping-rate of the heart. Psychologically induced heart rate is the increase in heart rate that cannot be attributed to physical activity, which is sometimes called "additional heart rate" (AHR; Carroll et al., 1987). When the actual heart rate lies above the heart rate that can be predicted by physical activity, this is called AHR (Carroll et al., 1987). Based on a method by Blix, Stromme, and Ursin (1974), who found that pilots had higher heart rates during challenging flight operations than would have been predicted by their oxygen consumption, AHR has traditionally been calculated through a regression function of heart rate and oxygen consumption. While some sort of physical task is executed, the participant's heart rate and oxygen consumption are measured. From these measurements, a regression equation with oxygen consumption as a predictor for heart rate is made. This regression function then serves as a predictor for the physical part of heart rate during a psychologically demanding task. The predicted heart rate is subtracted from the actual heart rate during the task, so that the additional heart rate is calculated. Different sorts of physical tasks, including isotonic exercise (Carroll, Turner, & Hellawell, 1986; Carroll, Turner, & Prasad, 1986), static exercise (Carroll et al., 1987), and upper and lower body dynamic exercise (Turner, Carroll, Hanson, & Sims, 1988) have already been researched as reference-exercises, and also different (difficulties of) psychological tasks were tested. Thayer, Van Doornen, Turner, and Building (1991) have also tested carbon dioxide production instead of oxygen consumption as a potential predictor for heart rate and found it to be appropriate. All these mentioned studies were executed in a laboratory, using controlled tasks with different levels of difficulty in order to simulate stressors. Wilhelm and Roth (1998) tried to bring this AHR-method into a more ambulatory context. Instead of oxygen consumption, they used minute ventilation and compared the AHRs of flight-phobics and non-phobics when entering a plane. Although their measurement-equipment was already wearable, it was far from real-life compatible, using multiple electrodes and sensors all over the body.

Another way to measure one's metabolic or muscle activity in order to calculate AHR is to measure body movement. Myrtek et al. (1988) developed a wearable device that measured acceleration and ECG. It compares heart rate and physical activity with heart rate and activity of the previous minutes in order to detect an increase in heart rate occurring without increase in physical activity. Their algorithm (Myrtek, Aschenbrenner, & Brügner, 2005) is based on the idea that an increase in heart rate that is not accompanied by an increase in physical activity must occur due to an emotional (psychological) event. This method was used in various studies (e.g. Myrtek, Aschenbrenner, & Brügner, 2005; Myrtek, Weber, Brügner, & Müller, 1996; Myrtek & Brügner, 1996). In a study by Myrtek, Weber, Brügner, & Müller (1996), it was found that chronically stressed students generally had a higher AHR than non-stressed students, and that this effect was especially true when they were at the university in contrast to being at home. Myrtek and Brügner (1996) also found that through AHR, emotional events could be detected, at least in the laboratory. AHR was higher for watching an erotic film compared to watching a comedy, both while being physically active and physically inactive (Myrtek & Brügner, 1996). For the field, the results concerning AHR were more ambiguous (Myrtek et al., 2005).

Because in the current study, the relationship between heart rate and psychological measures is researched, it is important to first filter out the part of the heart rate that correlates with physical activity. Both approaches, the one using oxygen consumption and the one by Myrtek, are not practical in real life. Measuring oxygen consumption is possible in the laboratory, but not practical in a study of daily life. The device developed by Myrtek et al. (1988) was wearable, but is not freely available for everyone. Also, two sensors are used in their approach, one on the chest and one on the hip. This is less comfortable and more handicapping in daily life than a presently often used wrist-sensor. For the current study, a new way of calculating a form of AHR is explored, which combines the two approaches and is possible with the sensors that are available in most wearables.

As with the traditional way of calculating AHR, the idea is to build a regression function to determine the part of the heart rate that can be predicted by physical activity. Normally, this regression is used to predict heart rate based on the participant's oxygen consumption. For the current study, Myrtek et al.'s (1988) idea to use physical activity instead of oxygen consumption to predict heart rate is used. This way, the problem with measuring oxygen consumption in real life is avoided. For measuring physical activity, an easier setup is used, since it can now already be measured with a wrist-worn accelerometer.

Physical activity. Since an accelerometer measuring 3-axis acceleration is built into most wearables already, its data is used to calculate the participants' physical activity. Accelerometers measure acceleration of the sensor as gravitational-force (g). 1g is equivalent to 9.81 m/s^2. In the current study, Euclidean Norm Minus One (ENMO) as described by Bakrania et al. (2016) is used to predict physical activity. It is an easy and reliable way of filtering out gravity, which always influences raw accelerometer data, so that pure motion data is received. ENMO is calculated through subtracting 1g from the magnitude of the vector of acceleration, which is calculated using the Euclidean Norm. All values that are negative after this subtraction are rounded up to 0g. The magnitude of acceleration calculated through ENMO is thus used as a predictive factor for heart rate.

RAHR. As individuals differ both in their heart rate and in their amount of physical activity, a regression line with magnitude of acceleration as a predictor for heart rate is fitted for each participant in the current study. Based on the fitted regression function, the physical part of the heart rate can be predicted through the magnitude of acceleration. A comparison of the estimated values to the observed values can provide an insight into the psychologically caused part of the heart rate. All positive residuals are thus possibly caused by psychological activity. The name of this new measure is "Residuals from Acceleration predicted Heart Rate" (RAHR). The next step in exploring this new measure is to correlate it to psychological measures, in this case stress and arousal.

Psychological Experience

The aim of the current study is to correlate physiological activity as measured through RAHR with psychological experience. As psychological experience is something subjective, it is best measured through the subjective measure of self-report. Different forms of subjective self-report exist (Conner & Feldman Barrett, 2012). On the one hand, there are retrospective self-report techniques which rely on memory-based reporting where participants need to remember and reproduce their experience. Here, the remembering self is examined. On the other hand, there are momentary self-report techniques, which examine the experiencing self. People are asked to describe what they are thinking or feeling in (near) real-time (Conner & Feldman Barrett, 2012). When those momentary self-reports take place in daily life and over a longer time, experience sampling is often used, which is also called ecological momentary assessment or ambulatory assessment (Conner & Feldman Barrett, 2012). Through the above-mentioned technological advances and mostly through the development of smartphones, experience sampling is easier to execute than ever. Participants can receive push messages to answer short questions which have been programmed into a mobile application. Those applications compile the acquired data so that they can be analysed easily (Conner & Feldman Barrett, 2012). To get a more complete picture and correlate both participants' momentary and retrospective self-report with RAHR, ratings of psychological experience in the current study are inquired for both the previous minute (momentary) and the previous two hours (retrospective). The two sorts of measures do not only differ between the type of self-report, but also concerning the recall period (i.e., one minute versus two hours). When certain stressful events need to be recalled by participants, the period over which these need to be recalled may affect the result (Cohen, Cimbolic, Armeli, & Hettler, 1998). The feeling of stress or arousal within one minute may be easier to determine for participants than the feeling of stress or arousal over two whole hours, since a lot can change within this timespan. There may thus be a difference between the reliance or validity of the two forms of measures, both due to the type of self-report and due to the recall period.

The two psychological variables that are correlated with RAHR in the current study are stress and arousal. As described above, stress level is a measure that users of wearables would like to have calculated for them in order to be able to actively relax themselves. Stress can be defined "as a state of high general arousal and negatively tuned but unspecific emotion, which appears as a consequence of stressors (i.e., stress-inducing stimuli or situations) acting upon individuals" (Boucsein, 2012, p. 381). This shows that stress is related to the 'fight-or-flight' system. When a potential danger is experienced, the body reacts to it by preparing to either fight the danger or run away from it (Segerstrom & Miller, 2004), a reaction experienced as stress. When stress becomes chronic, it can become a threat to humans' health (Sharma & Gedeon, 2012). Due to the fact that stress can potentially lead to health problems, and the fact that most wearable users' request to get feedback about their stress level, this psychological variable is included into the current study.

Concerning arousal, it can be said that it is not as extreme as stress and can be both negative or positive (Russell, 2009). In his circumplex model of affect, Russell (1980) presents a two-dimensional model featuring the dimensions pleasure-displeasure and activation-deactivation. Arousal is placed at high activation, but in the middle between pleasure and displeasure, showing that it is rather a general intensity of an emotion than its valence (Russell, 2009). Arousal is used as a psychological variable in the current study in addition to stress because it includes the intensity of both positive and negative feelings. Stress, in comparison, mostly includes negative feelings, although this may differ as some people may experience stress more positively than others. Still, Boucsein (2012) states that both negatively and positively experienced stress (i.e., distress and eustress) result in similar physiological responses. The current study does not differentiate between the two.

The current study

Additional heart rate seems to be a promising predictor for psychological activity. However, calculating it with oxygen consumption or a device built specifically for this is not practical for everyday use. For that reason, the present study explores a new approach for calculating additional heart rate called "Residuals from Acceleration predicted Heart Rate" (RAHR). Furthermore, the current study investigates the relationship between RAHR and momentary and retrospective self-reports of stress and arousal. The research question for the current study is thus: *How does the RAHR correlate to the participants' subjective experience of momentary and retrospective stress and arousal?*

Method

Participants

21 participants were recruited for the study through convenience sampling. Three participants did not complete the study due to errors with their E4 equipment and were removed from the data. The remaining 18 participants were aged 19 to 27 years (M = 20.83, SD = 1.86) and had either Dutch, German, or English as their native language. The participants were required to own a smartphone and be able to use a computer with an internet connection and a USB-port. The research was approved by the Ethics Committee of the faculty of behavioural sciences of the University of Twente.

Design

The study was executed in a longitudinal experience sampling study with a within-subject design. Both physiological data and psychological data were collected during a 7-day period. The physiological data were sampled passively and continuously through sensors. The psychological data were collected in a fixed time-based sampling method, asking the participants to actively answer four questions every two hours. Both momentary and retrospective self-report for stress and arousal had to be rated on a scale from 0 to 10.

Materials and measures

For the study's physiological measures, the participants were provided with Empatica E4 wristbands (Garbarino, Lai, Bender, Picard, & Tognetti, 2015). The wristband is equipped with different sorts of sensors. A photoplethysmograph (PPG) measures the wearer's blood volume pulse (BVP) with a sample rate of 64 Hz. From the BVP, the device can calculate the person's heart rate. The amount of physical activity of the wearer can be calculated from the built-in 3-axis accelerometer that measures the acceleration in all three directions at a 32 Hz sample rate. Measures that were not used for the present report, but also measured by the sensor, are electrodermal activity and skin temperature.

The participants received an E4 wristband in its container, a charging cradle, and a Micro USB to USB cable with which the wristband could be connected to a computer. They also received instructions on how to use the wristband. Through the software Empatica Manager, the participants uploaded the data collected over the respective day every night.

For the psychological measures used in the study, the participants installed the application "mQuest" on their smartphones. Through this application, surveys and questionnaires can be send to users in determined time intervals. Four questions were sent to them every two hours, "How intense were your emotions during the last two hours?", "How intense were your emotions during the last minute?", How much stress did you experience during the last two hours?", and "How much stress did you experience during the last minute?". The questions were rated on a scale from 0 (very low) to 10 (very high). Also, the Toronto Alexithymia Scale (TAS-20) was filled in once before and once after the study. The data from this scale were not used in the present report.

Procedure

Before the data gathering, the participants were briefed. They read and signed an informed consent and filled in the TAS-20, and received the E4 wristband as well as instructions on how to use it and how to upload the data every night. Also, an instruction for how to use the mQuest application was given after it was installed on the participants' phones.

During the data gathering, which took seven full days, the participants wore the E4 wristband all the time they were awake. During the seven days, the four self-report questions were prompted to them by the mQuest application every two hours. The participants uploaded the wristband data every night through their computer.

After seven days of data gathering, a debriefing was held for the participants. They gave back the E4 wristband and its accessories, and again filled in the TAS-20. Furthermore, a semi-structured exit interview was held, including questions about the burden of the experiment and the subjectivity of emotions and stress.

Data Analysis

The physiological data was obtained from the CSV files provided by the E4 sensor. Heart rate data was calculated from the inter-beat interval (IBI) data as estimated from the BVP by the E4 sensor. This was done by calculating 60 divided by the IBI values. As the raw 3-axis accelerometer data's unit of measurement was 1/64g, it was divided by 64 to scale it to +/- 2g. Physical activity was calculated from the three axes as the magnitude of the vector of acceleration, using the Euclidean Norm:

$$\text{Magnitude of acceleration} = \sqrt{x_i^2 + y_i^2 + z_i^2}$$

To filter out the effect of gravity on the accelerometer, the Euclidean Norm Minus One (ENMO) was calculated, where the magnitude of acceleration is subtracted by 1g. All values that come out of this calculation below zero are rounded up to zero. This way, the gravitational force should be excluded, leading to a measure for only physical activity.

The self-report data was obtained from the CSV files provided by the mQuest mobile application. All variables (HR, magnitude of acceleration, present stress, present arousal, past stress, past arousal) where provided with timestamps and joined together into a dataset based on the timestamps. The data were then averaged per minute. When a value was missing for at least one of the two physiological variables (HR or magnitude of acceleration), the whole row was removed from the dataset as a preparation for correlating the two variables with each other. From 91.500 rows of data, 68.829 rows remained after the removal of missing data. This high amount of missing values stems mainly from the missing values already present in the IBI file given by the E4 sensor, but also rows with only self-report data were removed, when no physiological data were measured for that time.

When plotting magnitude of acceleration against heart rate, only a small part of the acceleration data was above 0.2 g (Figure 2), specifically only about 0.2% of the data. As the confidence intervals in the curve drawn in the figure get bigger when the magnitude of acceleration becomes higher, it was chosen to remove the data points above 0.2 g acceleration and only work with the ones below that point. Through this, 119 of the total 68,829 rows were removed from the dataset. After this filtering (Figure 3), the regression line seems a little less steep towards x = 0.2, but in general it still looks similar to the line in Figure 2. This shows that not too much is different after the data points above 0.2 g are removed. Because the regression line in Figure 3 is nearly linear, a local linear correlation between heart rate and magnitude of acceleration is assumed for the part of the data that lies below 0.2g. This assumption is applied during the current study.

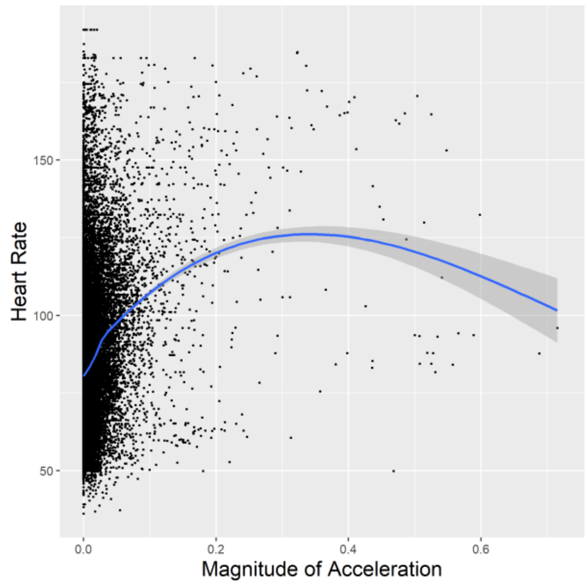

Figure 2. Magnitude of acceleration (g) plotted against heart rate (bpm).

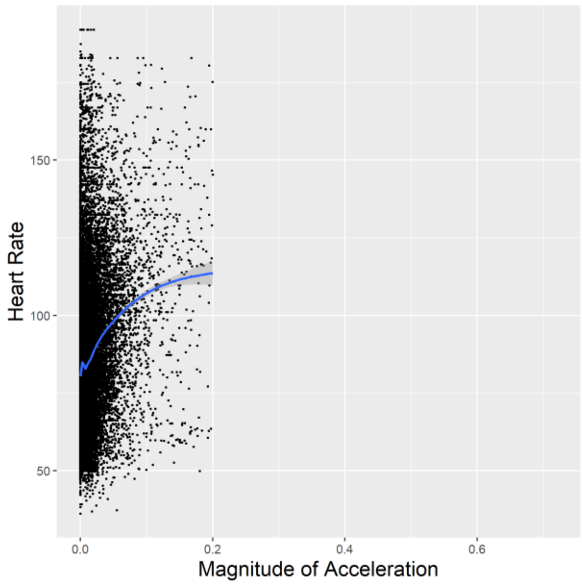

Figure 3. Magnitude of acceleration (g) plotted against heart rate (bpm) after removing acceleration above 0.2g.

In order to calculate RAHR, a mixed-effects regression model as described by Finch, Bolin, and Kelley (2014) was built by using the `lme4` package (Bates, Mächler, Bolker, & Walker, 2015) in the programme R (R Core Team, 2017), Version 3.4.2. A regression model was built with heart rate as dependent variable and magnitude of acceleration as predictor. A random effect for both intercept and slope was added for the different participants, because it seemed that both heart rate and physical activity were differing for the different participants (see section with descriptive statistics), so that including this into the model should provide a better fit. Also, because the estimated within-subject correlation of the random effect for the intercept and the random effect of the slope was found to be very low (-0.02), the easier model that does not allow for this correlation (i.e., in which two random-effects terms are specified) was chosen. In R, the final model looked like this:

```
HR.model <- lmer(HR ~ Acc + (1|Part) + (0+Acc|Part), data=D,
REML = FALSE)
```

For the calculation of the correlation of RAHR with the self-report variables, four models with RAHR as predictor variable and the self-report variables as dependent variable were constructed. Two were built for momentary self-report (i.e., stress and arousal during the last minute) and two for retrospective self-report (i.e., stress and arousal during the last two hours) data. For the momentary self-report, the data was correlated with the RAHR of the previous minute, and for the retrospective self-report, the data was correlated with the average RAHR of the previous two hours. The models that were fitted in R looked like this:

```
RAHRStrPres.model <- lmer(StressPres ~ RAHR + (RAHR|Part),
     data = DPres, REML=FALSE, control = lmerControl(optimizer
     = 'Nelder_Mead'))
RAHRAroPres.model <- lmer(ArousalPres ~ RAHR + (RAHR|Part),
     data = DPres, REML=FALSE, control = lmerControl(optimizer
     = 'Nelder_Mead'))
RAHRStrPas.model <- lmer(StressPas ~ RAHRmean +
     (RAHRmean|Part), data = DPas, REML = FALSE, control =
     lmerControl(optimizer = 'Nelder_Mead'))
RAHRAroPas.model <- lmer(ArousalPas ~ RAHRmean +
     (RAHRmean|Part), data = DPas, REML = FALSE, control =
     lmerControl(optimizer = 'Nelder_Mead'))
```

For these models, the Nelder-Mead optimization routine was used, because the BOBYQA optimization method which is used by the `lmer()`-function by default did not give any results for some of the model-fits.

Results

Descriptive Statistics

Physiological variables. The physiological variables that were measured for this study are heart rate, calculated from inter-beat intervals, and magnitude of acceleration, calculated with ENMO from 3-axis accelerometer data. In Table 1, a summary of the descriptive statistics for the two variables, averaged over all participants, can be found. The mean measured heart rate +/- one standard deviation reaches from 65.21 to 99.91 bpm, which lies inside the normal resting heart rate for adults (60 – 100 bpm; Laskowski, 2015). The mean for the magnitude of acceleration was found to be 0.01g, and there was not a lot of deviation from this. This shows that for the biggest part of the data the measured data points lie close to the mean showing not a lot of spread in acceleration. Because of the removal of the data with a magnitude of acceleration above 0.2 g, the maximum possible value lies at 0.2g.

Table 1. *Summary descriptive statistics of the averaged physiological variables*

Variable	N	Mean (SD)	Min	Max
Heart Rate	68710	82.56 (17.35)	36.22	191.99
Mag. Acc.	68710	0.01 (0.02)	0	0.2

When looking at the data from a between-subject perspective, variations can be found between participants in both heart rate (Figure 4) and physical activity (Figure 5). An example of two participants that differed substantially in their heart rate are 4 and 17. The upper quartile of participant 4 lies below the lower quartile of participant 17, which means that most of the time participant 17's heart rate lay above that of participant 4. Concerning the participants' physical activity, especially participant 4, 5, and 10 should be mentioned. The medians of their magnitude of acceleration are above 0g, which is not the case for the other participants. All in all, variation in the two physiological variables can be found between participants.

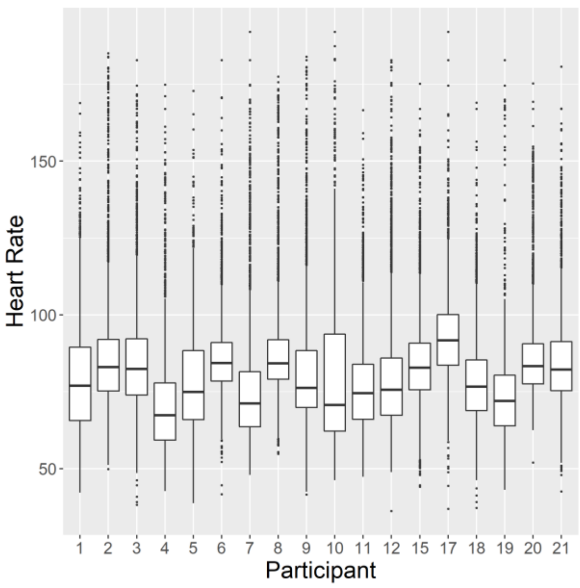

Figure 4. Heart rate (bpm) per participant

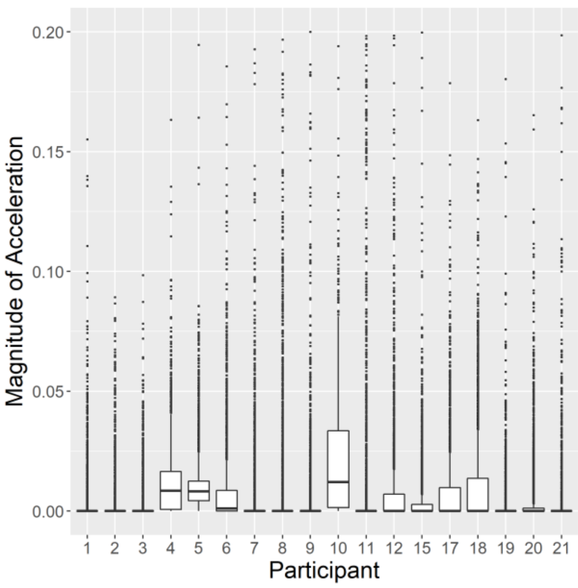

Figure 5. Magnitude of acceleration (g) per participant.

Self-Report variables. The four self-report variables that were measured during the current study are stress and arousal of the last minute (in the text also referred to as 'momentary', 'present', or 'now'), and stress and arousal for the last two hours (in the text also referred to as 'retrospective' or 'past'). The values reach from 0 (very low) to 10 (very high). As can be seen in Table 2, the values for retrospective arousal have the highest mean, but on average all data are on the low side.

Table 2. *Summary descriptive statistics of the averaged self-report variables*

Variable	Mean (SD)
Stress Now	1.86 (1.84)
Arousal Now	2.37 (1.74)
Stress Past	2.63 (1.99)
Arousal Past	3.34 (1.71)

N = 413, Range = 0 – 10

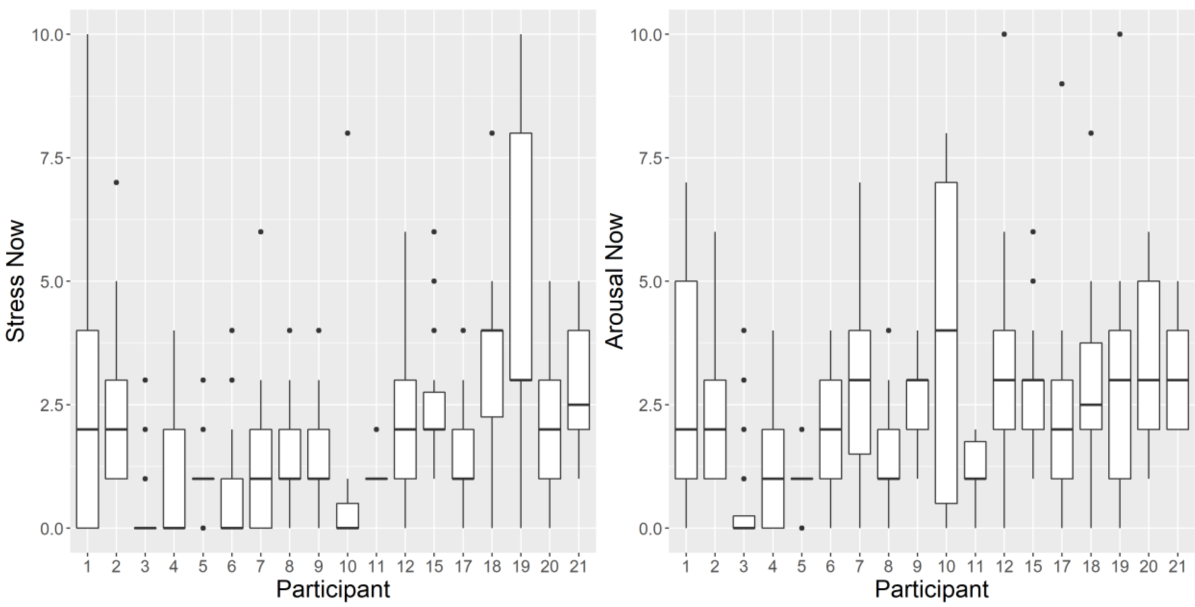

Figure 6. Answers to the momentary self-report variables per participant.

When looking at the between-subject data again, variations are also apparent in the self-report variables (Figure 6 and 7). In Figure 6 on the left, for example, the momentary subjective stress per participant is shown. Especially participant 19 stands out in this plot, because the upper quartile of his answers about momentary stress go up to 8, while it lies below 4 for most of the participants. In the right part of Figure 6, showing the momentary subjective arousal, especially participant 10 and his/her high spread between lower and upper quartile should be mentioned. In Figure 7 on the left, it can be seen that only participant 19 rated his past stress higher than 8 at least once. On the right side, participant 10 should be mentioned again, as his/her median is higher than all upper quartiles of the other participants.

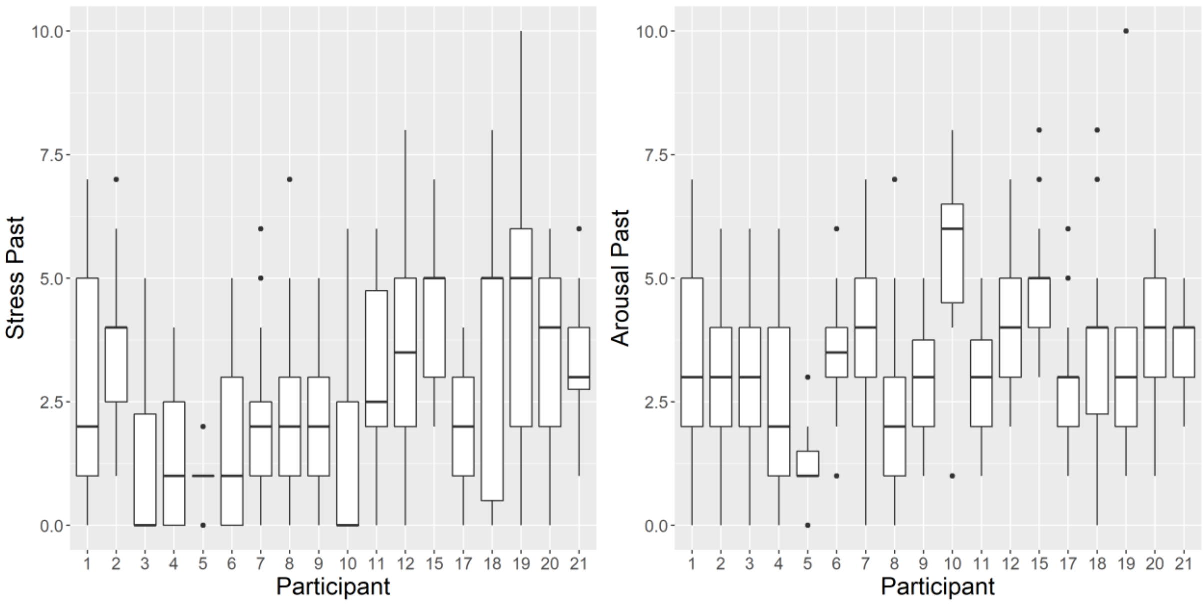

Figure 7. Answers to the retrospective self-report variables per participant.

Calculating RAHR

In order to calculate RAHR, a model fit was made with heart rate as an outcome, magnitude of acceleration as predictor, and random effects of participant on intercept and slope, which were uncorrelated. Below, the results for both fixed and random effects are explained.

Fixed effects. The fixed effect in the model was magnitude of acceleration as a predictor of heart rate. In order to make sure that magnitude of acceleration is indeed a predictor of heart rate, an ANOVA was executed, comparing the model including acceleration as a predictor with the model excluding acceleration as a predictor. It was found, that the model including acceleration was indeed a better fit for the data [$\chi^2(1) = 33.01$, $p < 0.001$], supporting its addition to the model. The test of statistical significance built into the R package `lmerTest` (Kuznetsova, Brockhoff, & Christensen, 2017) and based on the Satterthwaite approximation also shows that magnitude of acceleration is a significant predictor of heart rate in the fitted model [$t = 9.73$, $df = 20.63$, $p < 0.001$]. The direction of the prediction is positive, showing that the higher the value for magnitude of acceleration, the higher the heart rate.

When looking at only the fixed effects in the model, the pattern shown in Table 3 can be found. Here, it must always be remembered that these are averaged values. Because the model is built with participants as a random effect, these values may not be very meaningful, and an analysis of the random effects is still necessary. The intercept in Table 3 can be called the resting pulse, for it is the average heart rate at an acceleration of 0g, when a person is not physically active at all. The 95% confidence interval (CI) for the intercept does not leave a lot of space for deviation from the mean, ranging from 76.92 to 82.66 bpm. The slope that is given in Table 3 shows how much bpm the heart rate increases on average when magnitude of acceleration increases by 0.1g[1] - the higher the value for the slope, the bigger the increase. Here, a little more deviation from the mean is possible, so that the 95% CI leaves space for heart rate changes from 28.20 to 43.42 bpm once a person goes from inactive (0g) to a little active (0.1g). The regression line for this is the black line in Figure 8.

Table 3. Fixed effects, including 95% confidence intervals.

	2.5% CI	Mean (SD)	97.5% CI
Intercept	76.92	79.79 (1.39)	82.66
Slope 0.1g (Acc)	28.20	35.79 (3.68)	43.42

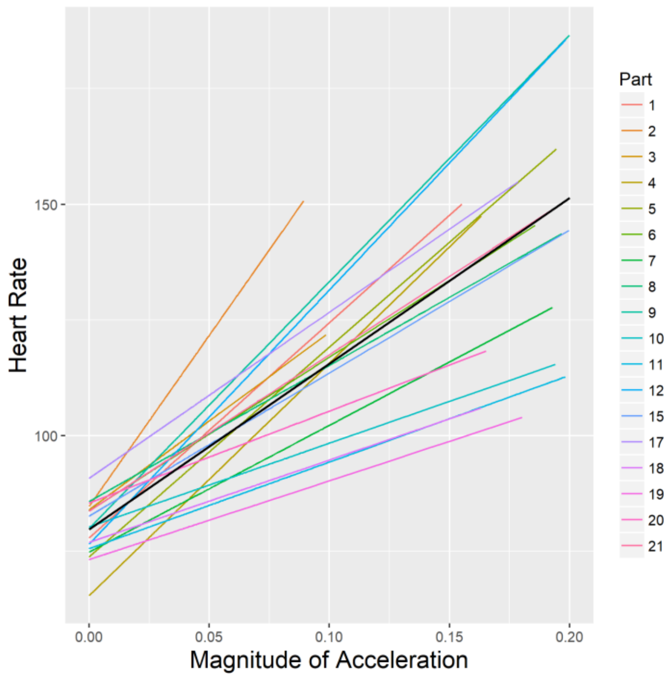

Figure 8. Linear regression lines for y = heart rate and x = magnitude of acceleration, plotted per participant. The black line shows the average regression line from the fixed effects data.

[1] Normally, the slope gives a value for how much the y-value (outcome variable, here heart rate) increases when the x-value (predictor, here magnitude of acceleration) grows by 1 (here 1g). Because a physical activity of 1g is not included in the data set, it was chosen to set 0.1g as a more meaningful increase in activity. For this, the original slope values were divided by 10. To get the original slope-values, the used values thus must be multiplied with 10.

Random effects. In Figure 8, the linear regression line for each participant based on the model with heart rate as outcome variable and magnitude of acceleration as predictor can be seen in comparison to the average regression line. The graph shows that the participants differ greatly in their relation between the two variables. Furthermore, an ANOVA shows that the mixed effects model (i.e., with both fixed and random effects) has a better fit for the data than a generalized linear model (i.e., with fixed effects, without random effects) [$\chi^2(2) = 7620.5$, $p < 0.001$]. This supports the decision of adding random effects to the model.

Table 4. *Coefficients of the mixed-effects model of heart rate and magnitude of acceleration per participant (Part.).*

Part.	Intercept	Slope 0.1g
1	77.96	46.47
2	84.80	73.85
3	84.06	38.31
4	65.50	50.19
5	73.88	45.29
6	83.79	33.20
7	74.88	27.39
8	85.81	29.41
9	80.00	53.28
10	80.32	18.08
11	75.62	18.72
12	76.62	54.90
15	82.67	30.90
17	90.84	35.86
18	76.97	17.81
19	73.26	17.03
20	85.48	19.86
21	83.75	33.77
Total	79.79	28.20

When looking at the coefficients of the model (Table 4), it can be seen that the participants' intercepts range from 65.50 to 90.84 (*M* = 79.79, *SD* = 6.03). The between-participant standard deviation of the intercept lies only at 5.87 (95% CI [4.37, 8.48]) (i.e., the variability of the intercept between participants is not very high), but when comparing the model with and without the random intercept effect through the `anova()` function in R, the model fit is significantly better for the model including the random intercept effect [$\chi^2(1) = 5612.2$, $p < 0.001$].

The slope also differs greatly between participants (Table 4). When looking at how much the participants' heart rate rises from no motion (0g) to some motion (0.1g), the value goes from minimally 17.03 to maximally 73.85 (*M* = 35.80, *SD* = 15.75). The slope's between-participant standard deviation is 15.45 (95% CI [11.39, 22.44]) (i.e., the variability of the slope between participants is quite high). Also, when comparing the model with and without the random slope effect through the `anova()` function, the model fit is again significantly better for the model including the random slope effect [$\chi^2(1) = 1042$, $p < 0.001$]. The big between-participants variance and the results from the ANOVAs show that it is appropriate to add participant as a random effect to the mixed-effects model, because more variance is explained by it.

The estimated within-subject correlation of the random effect for the intercept and the random effect of the slope was found to be very low (-0.02), which shows that a participant's high resting pulse did not predict a strong effect of physical activity on heart rate. When comparing the model including only one random-effects term for both effects with the model including one random-effects term for each effect through the `anova()` function, it was found that the two models do not differ significantly in their fit [$\chi^2(1) = 0.004$, $p = 0.9472$]. Because of these findings, the easier model excluding the correlation between random effect of participant on intercept and random effect of participant on slope was chosen.

RAHR. The residuals from this model (i.e. the difference between the expected values and the observed values) were taken as the RAHR values. All values lower than zero were replaced with a zero in the dataset. After this, RAHR was found to lie at maximally 116.32 bpm and minimally at 0 bpm (*M* = 5.67, *SD* = 11.56). Figure 9 shows the RAHR per participant. It can be seen that the median for each participant lies at zero, which means that more than half of all RAHR values are zero. Although the upper quartile is similar for most participants, especially participants 1 and 5 stand out with their higher upper quartiles and participants 8 and 20 with their lower upper quartiles. So, although not a lot of difference is apparent between the participants, there still is some variation.

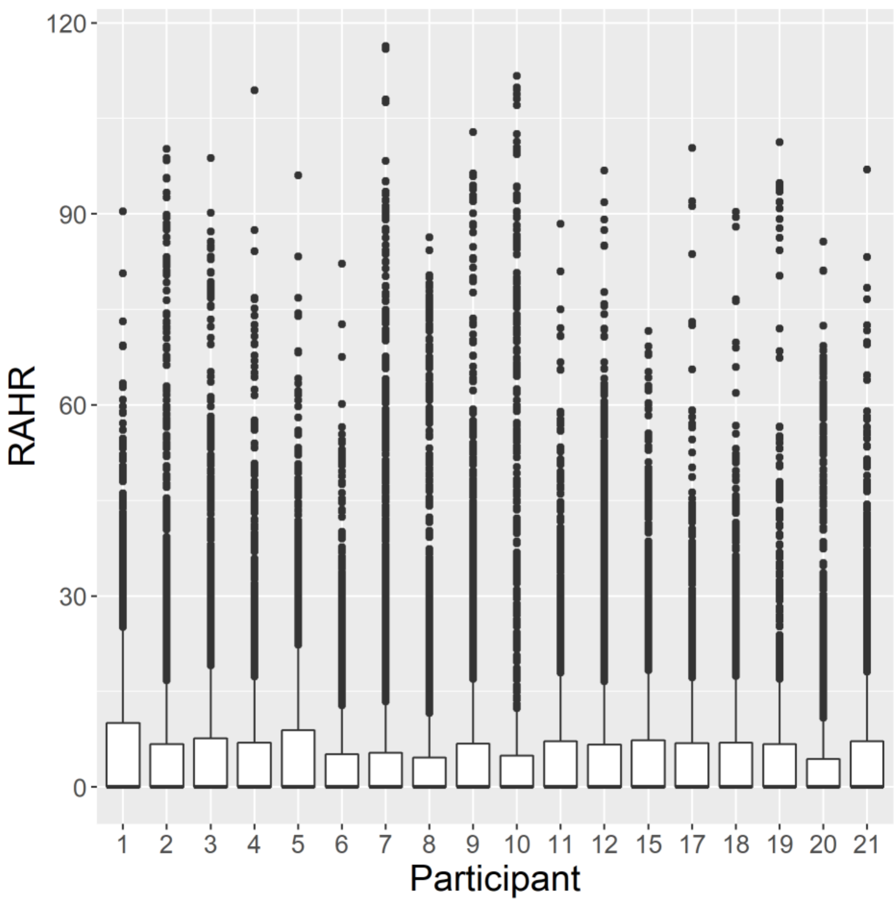

Figure 9. RAHR per participant

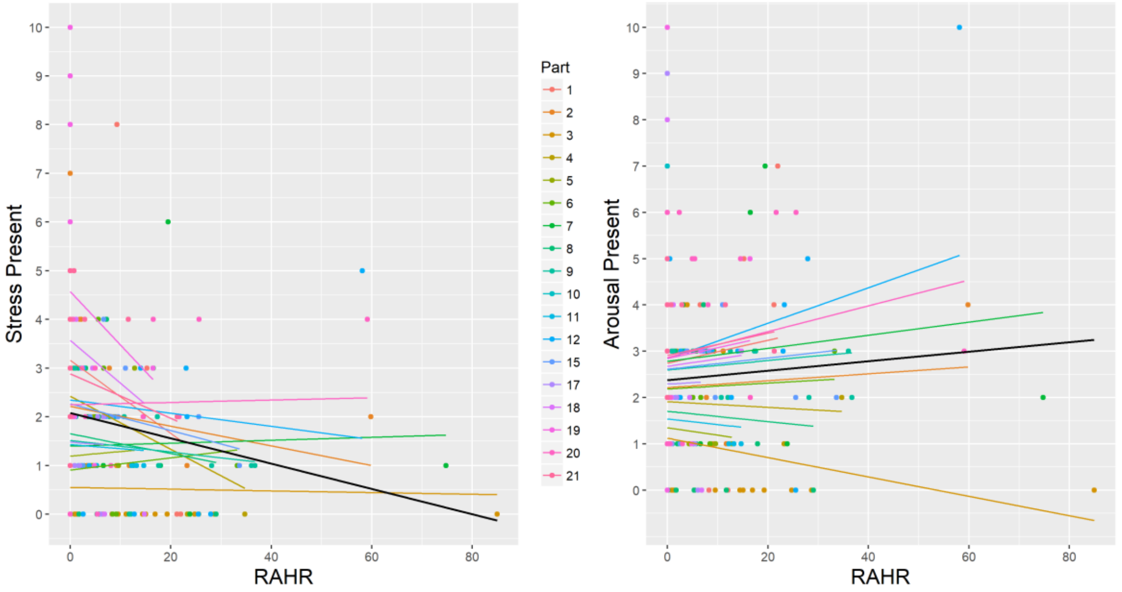

Figure 10. RAHR correlated with present stress and arousal.

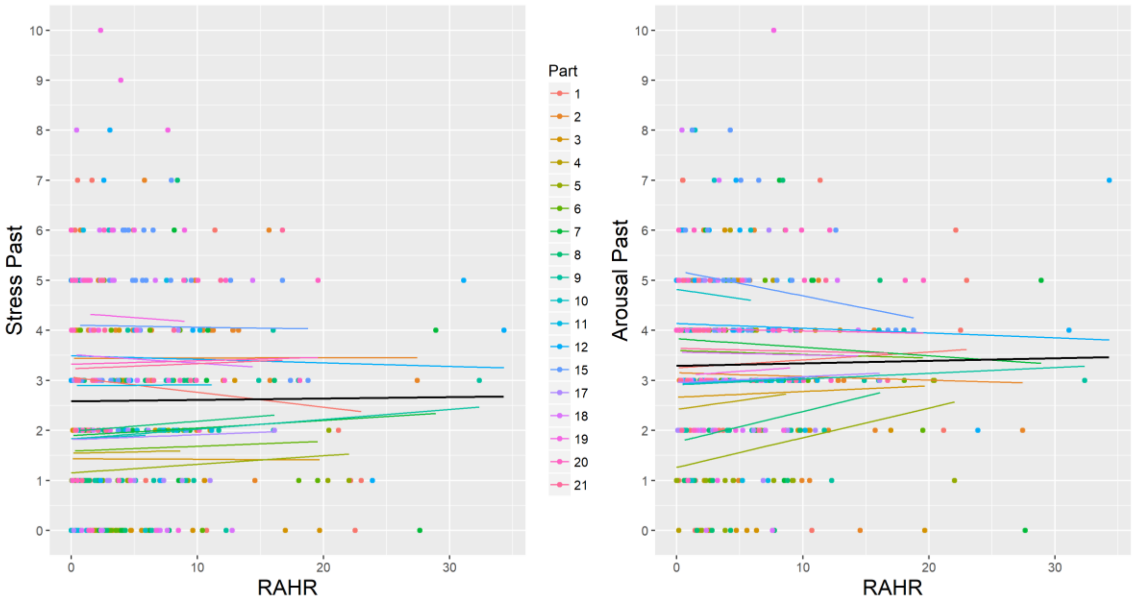

Figure 11. RAHR correlated with past stress and arousal.

Correlating RAHR and Self-Report Data

Based on the fitted models, with RAHR as predictor for the self-reported variables, the regression lines for each participant were drawn (Figure 10, 11). When looking at the left picture in Figure 10, it can be seen that RAHR and momentary self-reported stress seem to be correlated in the same direction for nearly all participants. The correlation is not significant [t = -1.87, df = 4.72, p = 0.124] and opposite to the expected correlation. Here it can be seen that the higher the self-reported stress, the lower the RAHR. The correlations of the other three self-report variables and RAHR are also not significant [(t = 0.98, df = 11.00, p = 0.347), (t =

0.17, df = 9.36, p = 0.869), (t = 0.33, df = 7.63, p = 0.754)]. In the graphs in Figure 10 and 11 it can be seen that individual participants seem to have a correlation between their RAHR and their self-reported data, but that these can be both positive and negative. No apparent pattern can be found for this.

The same approach was tried again including only the higher values of the self-reported variables (+1 SD above mean). The results found here were also non-significant. All in all, no apparent relation between self-reported data and RAHR calculated from heart rate and magnitude of acceleration could be found, neither on an inter-individual nor on an intra-individual level.

Discussion

The goal of this study was to explore the relationship between a new measure for the psychologically caused part of heart rate called "Residuals from Acceleration predicted Heart Rate" (RAHR) and self-reported momentary and retrospective stress and arousal data. With regard to the calculation of the RAHR it can be said that magnitude of acceleration and heart rate did indeed correlate with each other, so that the positive residuals of the data could be used as RAHR. In the next step, RAHR was correlated with different self-report data. Here, none of the correlations was found to be significant, indicating that there is no relationship could be found between RAHR and the subjective psychological data.

Concerning the calculation of RAHR, it was found that magnitude of acceleration indeed correlated with heart rate, so that the positive residuals of its regression line could be used as RAHR. This is in accordance with the studies that found that oxygen consumption (or minute ventilation, carbon dioxide production) predicted heart rate (Carroll, Phillips, & Balanos, 2009; Carroll et al., 1987; Carroll, Turner, & Hellawell, 1986; Carroll, Turner, & Prasad, 1986; Thayer et al., 1991; Turner et al., 1988; Wilhelm & Roth, 1998), although the methods used in those studies were quite different from the current study. Of course, the first big difference concerns the measures used: in the current study, physical activity instead of oxygen consumption was used as a predictor for heart rate. Although Wilhelm and Roth (1998) state that physical activity is not a good measure for metabolic demand, a clear correlation between heart rate and physical activity was found in the current study. The second difference concerns the data analysis. In the current study, one mixed-effects model was built for all data instead of one regression equation for each participant. By adding participants as random effects to the model, the individual differences were taken into account without the necessity to build 18 different regressions. The current method for analysis of the data thus seems to also be usable for bigger datasets, preventing the need to compute everything separately or losing track. Another difference to the previous studies concerns the design of the study. An approach was chosen that was easily executed in the field in comparison to the laboratory, were nearly all cited studies took place. The measurement of oxygen consumption is difficult outside of the laboratory, so that the previous studies were confined to that area. In the attempt by Wilhelm and Roth (1998) to execute an ambulatory study by swapping oxygen consumption for minute ventilation, still a lot of electrodes and sensors all over the body were necessary. In the current study, only one wrist-worn sensor was used. This way, participants could go on with their normal life without being disturbed or confined by too much equipment. By executing the study in the field, no artificial stressors were necessary. A fourth difference to the previous studies was the length of the current study. While all cited studies took place during a few minutes, hours, or at the most a day, the current study ran seven whole days. This way, a lot more data could be collected over a lot of different real-life situations. Although the current study was so different from the previous studies, a correlation was found based on which RAHR could be calculated. Here, it is especially striking that the correlation was different for all participants, showing the necessity of person-specific approaches when calculating this sort of data.

Contrary to the hope to validate the RAHR against psychological data, no significant relationship was found between the psychological and physiological variables. There could be multiple reasons for not finding a relationship between the variables. One explanation is that the newly calculated measure RAHR might not be fitting to be correlated with subjective psychological data. Due to some limitations with the measure, it may be the case that RAHR may not be an appropriate measure to predict psychological variables. When comparing the present study's way of calculating RAHR to previous approaches of calculating AHR, one difference stands out: in other studies, the regression line was built based on a baseline measurement, in which the participants were physically active. In the current study, the regression line was made for the whole time of measurement, so that it is not excluded that there is already a part of the psychologically explained heart rate in this regression line. This part may be different for every participant and bias the results. In a potential future study, it is suggested to try the new calculation with a baseline approach.

Another possible explanation for not finding significant results is that RAHR is appropriate to predict psychological data, but not the four variables measured during this study. It could, for example, be usable for the prediction of specific emotions like fear (e.g., flight-phobia, Wilhelm and Roth (1998)) or anger (which seems to have a big influence on cardiovascular activity (Suchday, Carter, Ewart, Larkin, & Desiderato, 2004)), or cognitive workload (as in the original AHR study by Blix et al. (1974), which examined pilot workload).

When looking at the different categories of relation between psychological and physiological events by Cacioppo et al. (2017), the relation between the psychological and physiological elements measured in the current study is definitely not found to be an easy one-to-one relation. More research is necessary to determine which of the other four categories may be true for these variables. Based on the findings in the current study, it looks like there is no relation at all, which would be in accordance with the study by Evers et al. (2014). Here, the authors presented a framework showing that a relation between responses only takes place within one system (either physiological reactions or psychological experience), but not across those systems. In the context of the current study, this means that the non-existent correlation of the variables might be explained by the assumption that physiological reactions (i.e., heart rate) are part of the automatic system, and experience of psychological states (i.e., stress and arousal) are part of the reflective system. There should, however, be a correlation between subjective stress and arousal in the current study because both are part of the reflective system. In a follow-up analysis, these correlations between the pairs present stress – present arousal, and past stress – past arousal could indeed be found, providing further support for Evers et al.'s (2014) claims that there might be no response coherence between the two domains but only within them. Other authors have also stated that the relation between physiological responses of the body and psychological experience of the mind is more complex than one would think (e.g., Fairclough, 2009; Feldman Barrett, 2006).

Strengths and Limitations

When looking at the current study, several strengths can be seen. The biggest strength is that a new approach for calculating additional heart rate has been tested. It can easily be implemented in most wearables, because it only needs heart rate and accelerometer data, which are both often measured in common sensors. Even though it was not possible to correlate the measure with the self-reported data in the current study, this does not mean that it is completely worthless. As people may not know about how their body reacts to psychological stress or arousal, RAHR may be able to provide them with some insight to this. Most wearables only give normal heart rate, in which the physically caused part is also included. RAHR could help people in learning more about their body, because direct feedback about bodily functions is possible. This might then even lead to a higher correlation of physiological and psychological variables (Van Dijk, Westerink, Beute, & Ijsselsteijn, 2015). Another strength is the big amount of data that could be gathered for the study. More than 65.000 points of aggregated physiological data and more than 400 points of self-report data were accessible. Due to this amount of data, the positive correlation of magnitude of acceleration and heart rate was found to be significant. Also, for each individual a lot of data points were available so that individual differences could be exposed, although the correlation was positive for every participant. The field-test nature is also a strength, as it is more promising than laboratory studies concerning ecological validity.

Of course, the study also has its limitations. Although there were already a lot of data points, the participants only took part in the study for seven days. People may differ from situation to situation, and executing a study in daily life over a longer period gives the opportunity for a bigger variance of different situations. If, for example, during the measured week some of the participants had an exam week, this would probably have influenced their results. This shows that the field-test nature also bears its limitations. No control and no experimental manipulations are possible in daily life.

In this study, the “correctness” of the RAHR can only be inferred from the fact that magnitude of acceleration and heart rate significantly correlate with each other. For a future study, it is important to compare the new way of calculating RAHR with the other ways of calculating AHR that are already used (e.g., with the regression line of heart rate and oxygen consumption) in order to validate it further.

The experience sampling part of the study could also be improved. The terms “stress” and “arousal” seem quite general and may not catch the emotions that RAHR may be able to find. For more clarity for both participants and researchers, more specific data on what emotions people feel at a moment could be inquired. What could be added to the experience sampling part of the study in a future research is that the participants are not only asked about their feelings, but also about the kind of situation they are in at that moment. This would make an exploration of even more correlations possible and may lead to an identification of situations in which people feel certain emotions in their daily life, and maybe even an identification of the situations in which RAHR and psychological measures are more related with each other. Also, a little bit more control is added to the quite uncontrollable nature of a field test.

Conclusion

The current study explored a new approach for calculating the psychologically caused part of heart rate (RAHR) and examined the relationship between this newly calculated variable and both momentary and retrospective self-report. Although the results have not all been as expected, they can be taken as a starting point for further research concerning for example wearables, response coherence, RAHR, or daily-life data collection.

References

Bakrania, K., Yates, T., Rowlands, A. V., Esliger, D. W., Bunnewell, S., Sanders, J., Davies, M., Khunti, K., & Edwardson, C. L. (2016). Intensity thresholds on raw acceleration data: Euclidean norm minus one (ENMO) and mean amplitude deviation (MAD) approaches. PLoS ONE, 11(10), 1–16. https://doi.org/10.1371/journal.pone.0164045

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/doi:10.18637/jss.v067.i01

Blix, A. S., Stromme, S. B., & Ursin, H. (1974). Additional heart rate - an indicator of psychological activation. *Aerospace Medicine*, *45*(11), 1219–1222.

Boucsein, W. (2012). *Electrodermal Activity* (2nd ed.). New York, NY, USA: Springer Science + Buisness Media. https://doi.org/10.1007/978-1-4614-1126-0

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (2017). Strong inference in psychophysiological science. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (4th ed., pp. 3–15). Cambridge, UK: Cambridge University Press.

Caddy, B. (2017). Stress-beating tech to keep you sane. Retrieved October 22, 2017, from https://www.wareable.com/wearable-tech/stress-beating-tech-to-keep-you-sane

Carroll, D., Phillips, A. C., & Balanos, G. M. (2009). Metabolically exaggerated cardiac reactions to acute psychological stress revisited. *Psychophysiology*, *46*(2), 270–275. https://doi.org/10.1111/j.1469-8986.2008.00762.x

Carroll, D., Turner, J. R., & Hellawell, J. C. (1986). Heart rate and oxygen consumption during active psychological challenge: The effects of level of difficulty. *Psychophysiology*, *23*(2), 174–181.

Carroll, D., Turner, J. R., & Prasad, R. (1986). The effects of level of difficulty of mental arithmetic challenge on heart rate and oxygen consumption. *International Journal of Psychophysiology*, *4*, 167–173. https://doi.org/10.1016/0167-8760(86)90012-7

Carroll, D., Turner, J. R., & Rogers, S. (1987). Heart rate and oxygen consumption during mental arithmetic, a video game, and graded static exercise. *Psychophysiology*, *24*(1), 112–118.

Cohen, L. H., Cimbolic, K., Armeli, S. R., & Hettler, T. R. (1998). Quantitative assessment of thriving. *Journal of Social Issues*, *54*(2), 323–335. https://doi.org/10.1111/j.1540-4560.1998.tb01221.x

Conner, T. S., & Feldman Barrett, L. (2012). Trends in Ambulatory Self-Report.

Psychosomatic Medicine, *74*(4), 327–337. https://doi.org/10.1097/PSY.0b013e3182546f18

Evers, C., Hopp, H., Gross, J. J., Fischer, A. H., Manstead, A. S. R., & Mauss, I. B. (2014). Emotion response coherence: A dual-process perspective. *Biological Psychology*, *98*(1), 43–49. https://doi.org/10.1016/j.biopsycho.2013.11.003

Fairclough, S. H. (2009). Fundamentals of physiological computing. *Interacting with Computers*, *21*(1–2), 133–145. https://doi.org/10.1016/j.intcom.2008.10.011

Feldman Barrett, L. (2006). Are Emotions Natural Kinds? *Perspectives on Psychological Science*, *1*(1), 28–58. https://doi.org/10.1111/j.1745-6916.2006.00003.x

Finch, W. H., Bolin, J. E., & Kelley, K. (2014). Multilevel Modeling Using R. *CRC Statistics in the Social and Behavioral Sciences*. Retrieved from http://www.crcnetbase.com/isbn/978-1-4665-1586-4

Garbarino, M., Lai, M., Bender, D., Picard, R. W., & Tognetti, S. (2015). Empatica E3 - A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. *Proceedings of the 2014 4th International Conference on Wireless Mobile Communication and Healthcare (Mobihealth)*, 39–42. https://doi.org/10.1109/MOBIHEALTH.2014.7015904

IDC. (2017). Worldwide Wearables Market to Nearly Double by 2021, According to IDC. Retrieved October 22, 2017, from https://www.idc.com/getdoc.jsp?containerId=prUS42818517

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology*, *84*(3), 394–421. https://doi.org/10.1016/j.biopsycho.2010.03.010

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. Retrieved from 10.18637/jss.v082.i13

Laskowski, E. R. (2015). What's a normal resting heart rate? Retrieved December 3, 2017, from https://www.mayoclinic.org/healthy-lifestyle/fitness/expert-answers/heart-rate/faq-20057979

Maslakovic, M. (2017). Ten stress busting wearables to help you chill. Retrieved December 10, 2017, from http://gadgetsandwearables.com/2017/11/08/stress-relaxation/

Myrtek, M., Aschenbrenner, E., & Brügner, G. (2005). Emotions in everyday life: An ambulatory monitoring study with female students. *Biological Psychology*, *68*(3), 237–255. https://doi.org/10.1016/j.biopsycho.2004.06.001

Myrtek, M., & Brügner, G. (1996). Perception of emotions in everyday life: Studies with patients and normals. *Biological Psychology*, *42*(1–2), 147–164. https://doi.org/10.1016/0301-0511(95)05152-X

Myrtek, M., Brügner, G., Fichtler, A., König, K., Müller, W., Foerster, F., & Höppner, V. (1988). Detection of emotionally induced ECG changes and their behavioural correlates: a new method for ambulatory monitoring. *European Heart Journal*, *9 Suppl N*, 55–60.

Myrtek, M., Weber, D., Brügner, G., & Müller, W. (1996). Occupational stress and strain of female students: results of physiological, behavioral, and psychological monitoring. *Biological Psychology*, *42*(3), 379–391. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8652754

R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition and Emotion*, *23*(7), 1259–1283. https://doi.org/10.1080/02699930902809375

Sandulescu, V., Andrews, S., Ellis, D., Bellotto, N., & Mozos, O. M. (2015). Stress Detection Using Wearable Physiological Sensors. In J. M. F. Vincente, J. R. Álvarez-sánchez, F. De la Paz López, F. J. Toledo-Moreo, & H. Adeli (Eds.), *Artificial Computation in Biology and Medicine, International Work-Conference on the Interplay Between Natural and Artificial Computation, Part I* (pp. 526–532). Elche, Spain: Springer. https://doi.org/10.1007/978-3-319-18914-7_55

Segerstrom, S. C., & Miller, G. E. (2004). Psychological Stress and the Human Immune System: A Meta-Analytic Study of 30 Years of Inquiry. *Psychological Bulletin*, *130*(4), 601–630. https://doi.org/10.1037/0033-2909.130.4.601

Sharma, N., & Gedeon, T. (2012). Objective measures, sensors and computational techniques for stress recognition and classification: A survey. *Computer Methods and Programs in Biomedicine*, *108*(3), 1287–1301. https://doi.org/10.1016/j.cmpb.2012.07.003

Suchday, S., Carter, M. M., Ewart, C. K., Larkin, K. T., & Desiderato, O. (2004). Anger cognitions and cardiovascular recovery following provocation. *Journal of Behavioral Medicine*, *27*(4), 319–341. https://doi.org/10.1023/B:JOBM.0000042408.80551.e1

Thayer, J. F., Van Doornen, L. J. P., Turner, J. R., & Building, M. (1991). Calculation of additional heart rates using oxygen consumption and carbon dioxide production: A comparative analysis or to. *Behavior Research Methods. Instruments. & Computers*, *23*(1), 2–4.

Turner, J. R., Carroll, D., Hanson, J., & Sims, J. (1988). A comparison of additional heart rates during active psychological challenge calculated from upper body and lower body dynamic exercise. *Psychophysiology*, *25*(2), 209–216.

Valencell. (2016). National Wearables Survey Reveals Accuracy is Top Priority Among Consumers; Lack of Continually Interesting Insights Among Top Reasons for Discontinued Use. Retrieved October 22, 2017, from https://valencell.com/press/2016/06/national-wearables-survey-reveals-accuracy-is-top-priority-among-consumers-lack-of-continually-interesting-insights-among-top-reasons-for-discontinued-use/

Van Dijk, E. T., Westerink, J. H. D. M., Beute, F., & Ijsselsteijn, W. A. (2015). In Sync: The Effect of Physiology Feedback on the Match between Heart Rate and Self-Reported Stress. *BioMed Research International*, *2015*. https://doi.org/10.1155/2015/134606

Wijsman, J., Grundlehner, B., Liu, H., Hermens, H., & Penders, J. (2011). Towards mental stress detection using wearable physiological sensors. In *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 1798–1801). Boston, MA, USA: IEEE. https://doi.org/10.1109/IEMBS.2011.6090512

Wilhelm, F. H., & Roth, W. T. (1998). Using minute ventilation for ambulatory estimation of additional heart rate. *Biological Psychology*, *49*(1–2), 137–150. https://doi.org/10.1016/S0301-0511(98)00032-5

Zanstra, Y. J., & Johnston, D. W. (2011). Cardiovascular reactivity in real life settings: Measurement, mechanisms and meaning. *Biological Psychology*, *86*(2), 98–105. https://doi.org/10.1016/j.biopsycho.2010.05.002