Measuring Stress: A Cross-instrumental Validity and Reproducibility Study

Theres Patzelt

S3973832

Department of Psychology, University of Groningen

PSB3E-BT15: Bachelor Thesis

Group 45

Supervisor: dr. Mark Span

Second evaluator: (prof.) (dr(s).) First name Last name

In collaboration with: Nienke Buist, Harmien Tamsma, Rover Willemars & Lisanne Zondag

June 20, 2022

A thesis is an aptitude test for students. The approval of the thesis is proof that the student has sufficient research and reporting skills to graduate, but does not guarantee the quality of the research and the results of the research as such, and the thesis is therefore not necessarily suitable to be used as an academic source to refer to. If you would like to know more about the research discussed in this thesis and any publications based on it, to which you could refer, please contact the supervisor mentioned.

Abstract

Background: Electrocardiography (ECG) is commonly used in psychological research and clinical practice; however, ECG machines are impracticable. This study cross-validates two wireless alternatives (Empatica E4 and polar H10) and assesses whether they agree with the golden standard. The monitors are normed to the REFA. To assess the validity of these devices, participants performed different tasks. **Methods:** An within-subject experiment was conducted. Participants (N= 28) performed a resting, physical and mental task (Stroop task and emotional Stroop) while being attached to the three devices. **Results:** The polar was highly agreeable and correlated with the REFA. The devices did not find a difference between the Stroop tasks. The devices found a difference between the physical and resting task. **Limitations:** The Empatica E4 did not collect enough data points to be accurately compared. **Conclusion**: The polar is highly agreeable with the REFA and can be used as valid alternative.

Keywords: Electrocardiogram (ECG), Polar H10, Empatica E4, Stroop task, cross-validity, psycho-motor vigilance

Measuring Stress: A Cross-instrumental Validity and Reproducibility Study

Introduction

Heart rate monitoring is a standard assessment in research and clinical practice. However, the commonly used ambulatory ECG machine is less practicable in the actual research setting. It needs to be connected to a box; therefore, it is often avoided in research (Millstein et al., 2020). In contrast, wireless heart monitors are becoming increasingly popular. Even though typically developed to assess athletes, whether they are reliable tools in research becomes more relevant among researchers. Wireless devices are more compact and practical, less expensive, and have fewer restrictions on natural movements such as walking (Millstein et al., 2020). Compared to ambulant ECG devices, they pose a valuable alternative. However, their validity and reliability need to be established. Wireless ECG devices are yet to be successfully integrated into clinical and research practice. This report will present the basics of an Electrocardiogram, introduce the devices assessed in the study, and evaluate whether they can detect differences in participants physiology while performing physical and mental tasks.

Electrocardiogram (ECG)

Specialized cells in the heart send electrical impulses that produce myocardial, i.e. muscular contraction (Bayés de Luna et al., 2022). An ECG is a continuous recording of cardiac electrical activity, i.e. electrical impulses. Each cardiac cycle consists of a P-wave, a QRS complex and T-wave. Each cardiac cycle follows the order mentioned earlier in a sinus rhythm, i.e. P-QRS-T (Fig. 1). (de Luna et al., 2019).

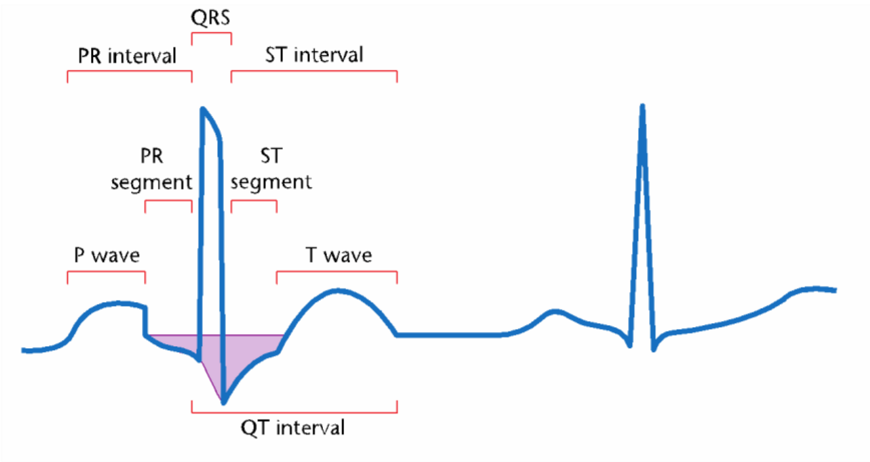


Figure 1 ECG morphology recorded in left ventricle (de Luna et al., n.d.)

P-wave

The p-wave is also called the atrial depolarization wave, i.e. change in electrical charge with a general height not exceeding 2.5 mV and width of 0.1 seconds (de Luna et al., 2019.).

PR interval and segment

The PR interval typically lasts from 0.12 to 0.2 seconds. It is the distance from the beginning of the p-wave until the beginning of the QRS-complex. The PR segment is a descending segment that follows from the end of the p-wave and lasts until the beginning of the QRS-complex (de Luna et al., 2019). The downsloping of the PR segment is explained by a sympathetic overdrive (de Luna et al., 2019), i.e. the Sympathetic nervous system is overactive compared to the parasympathetic system.

QRS complex

QRS, also called ventricular depolarization wave, begins at three different sites and occurs in three consecutive vectors. The QRS is a triphasic complex, i.e. it consists of three waves, the Q-wave, R-wave and S-wave. The Q-wave is the initial deflection; the Q-wave is typically broader and deeper than the R- and S-wave. If the R- and S-wave amplitude is large, they and written in capital letters. The overall length of the QRS-complex should not exceed 0.095 seconds (de Luna et al., 2019)

QT-interval

The QT-interval is the sum of depolarization and repolarization, i.e. QRS-complex and the ST-segment and T-wave (de Luna et al., 2019).

ST-segment and T-wave

The ST segment connects the QRS complex with the T-wave. The T-wave is called the ventricular repolarization or recovery wave, meaning the electrical charge returns to its original negative state (Bayés de Luna et al., 2022). follows the ST segment. The T-wave has a slower ascending slope than a descending slope, i.e., it takes longer to reach its peak than to go down again (de Luna et al., 2019).

Units of Heartbeat: Heart rate, Heart rate variability and Interbeat Intervals

Heart rate monitoring is a standard assessment in research. Heart rate (HR) describes the number of heartbeats per minute, and Heart rate variability (HRV), the fluctuation of time between two adjacent heartbeats (Chattopadhyay & Das, 2021). HR and HRV both originate from the automatic nervous system (ANS) to adapt to external, i.e. environmental, and internal, i.e. psychological challenges. This flexibility allows the heart to fluctuate and rapidly cope with a changing environment, explaining why healthy hearts are complex and respond non-linear (Shaffer & Ginsberg, 2017).HRV is understood in terms of Interbeat Intervals (IBI) (Prokhorov et al., 2021) as the period between two successive heartbeats. IBIs are measured in seconds or ms. IBIs vary depending on the HR. The faster the HR, the less time in-between the successive heartbeats and, therefore, the shorter the IBI and a lower HRV. In contrast, the lower the HR, the more time in-between two successive hearts beats, resulting in a longer IBI and a raised HRV (Shaffer & Ginsberg, 2017).

ECG devices

The Empatica E4 wristband (EE4) and the Polar H10 may be promising alternatives. Both are compact and wireless, whereas an ECG typically has electrodes connected to a box. These alternative measurements are promising, non-invasive, inexpensive research tools with comparable quality.

Empatica E4

The Empatica E4 (EE4) is a wristband and, therefore, less invasive. According to Schuurmans et al. (2020), the EE4 provides accurate heart rate variability measurements compared to the golden standard ECG machine (Schuurmans et al., 2020). The EE4 estimates IBIs based on a photoplethysmograph that sends and then detects the intensity in which the light is refracted (Milstein et al., 2020). However, the EE4 has some drawbacks. The EE4 is sensitive to movement and pressur. The inbuilt algorithm removes beats if the difference between two consecutive beats is not in line with the intervals of the other IBIs (Milstein et al., 2020). This will ultimately lead to missing data (Schuurmans et al., 2020). Nevertheless, the EE4 is a promising research tool to examine. It's easy administration and if accurat for HR a valuable research tool.

The polar H10

The polar H10 is considered a golden standard measuring device (Chattopadhyay & Das, 2021). According to the polar H10 website, the polar band measures HR and Heart rhythm, just like an ambulant ECG machine. It needs to be strapped around the chest where different sizes are available (polar.com, n.d.); therefore, it can be used by various body sizes. The H10 is strapped around a person's chest to have contact with the skin (Chattopadhyay & Das, 2021). The polar band has several advantages. Firstly, it is placed closed to the hearta dn less sensitive to movements. According to the website (polar.com, n.d.), the wrist band tolerates sweat, harsh conditions, electric disturbances, and motion artefacts. Lastly, the Polar H10 is inexpensive and can be purchased for 80 € (polar.com, n.d.). Compared to the clinical

ECG, a chest belt is less invasive and more comfortable. Comfortability is particularly important for a clinical population or if participants must complete long or complex tasks during an experiment. Although developed for athletes, the polar website (polar.com, n.d.) states that it is widely used in academic and industrial research (Kunkel et al., 2021). The affordability of the Polar H10 is a tremendous advantage in research practice, as it can be used in higher quantities. The polar H10 is a promising wireless alternative to ambulatory ECG machines due to its accuracy, low price, and robustness for movement. However, it may be uncomfortable to wear compared to the EE4.

Stress and Heart rate Variability

Stress is a bodily reaction reflected in physiological measures such as HR. HRV may be a physiological indicator of whether an individual is relaxed or stressed. Generally, there is no universally accepted definition for stress (Kim et al., 2018). This study defines stress as a bodily reaction to any stimulus that may disrupt the mental or physical balance. The main goal is to maintain homeostasis by activating the autonomic nervous system (ANS). The ANS consists of the parasympathetic nervous system (PNS) and sympathetic nervous system (SNS). With the SNS having an activating, energizing effect and the PNS a resting effect. The ANS prepares the body for a possibly stressful situation (Chattopadhyay & Das, 2021). Any small or more significant deviance from normal may be described as stress and can cause physical and psychological impairments (Pinel, 2020). A potential biomarker for stress may be Heart rate variability as the ANS influences cardiovascular muscles that, in turn, impact HR and HRV (Chattopadhyay & Das, 2021). The fluctuation in the length and frequency of heartbeat intervals indicates a person's general health and suggests momentary states (Kim et al., 2018). Awareness of a person's momentary state is relevant in a clinical and research setting. Therefore, HRV is a valuable biomarker to indicate an active sympathetic nervous system.

Physical and Mental activity and HR

Typically, deviations in HR are used to assess physical abilities, i.e. how the body performs while moving; however, HR is also a good indicator of mental stress. Physical activity is any bodily movement ranging from moderate walking to marathon running. Physical activity requires energy expenditure and motor behaviour, thereby provoking cardiovascular activity (Satish et al., 2015; Stamou et al., 2020), resulting in increased bodily functions such as HR. Therefore, HR should increase in response to engaging in physical activity.

Less direct but still relevant is the influence of mental activity on HR. Compared to physical activity, the reactivity of HR to mental activity or stress has more significant individual differences within individuals. However, HRV is still a proper measurement to detect possible stress in individuals. The brain innervates the Autonomic Nervous system in response to any stimuli. In a stressful event, the SNS may accelerate HR. At the same time, the parasympathetic nervous system engages in measures to balance bodily reactions to normal levels. The extent to which individuals may experience stress may differ but may be acute or chronic (Taelman et al., 2009). Thereby possibly affecting HR when being emotionally challenge.

STROOP and psychomotor task

Various psychological tests are available to accelerate mental activity (i.e. stress). The Stroop colour-naming task consists of congruent and non-congruent sub-parts. In the non-congruent part, colour and words are conflicting, increasing difficulty and ANS (i.e. stress) levels observed in the HR. While evaluating ECG, Karthikeyan et al. (Satish et al.,2015) observed significant variation and increased HR during the performance of the Stroop task compared to a relaxing control task (Karthikeyan et al., 2014; Vazan et al., 2017). The Stroop test as a mental stress condition has been found to activate the SNS resulting in HR changes, and is, therefore, an appropriate assessment of short duration mental stress (Satish et al., 2015)

A psychomotor vigilance task (PVT) assesses attention; however, it also can be used to induce mental stress. If it disrupts a different task, a startle reaction can be observed. The PVT aims to assess attention and reaction time. A visual or auditory stimulus occurs randomly and requires attention (Luque-Casado et al., 2013). The PVT may be used singular but can also be used during another test, such as the Stroop test an unpleasant stimulus disrupts the initial task. The unpleasant and sudden stimuli influence the activity of the ANS (Chua et al., 2012). They are, therefore, a practical test to induce change in HR.

The Present Study

This study aims to test whether the wireless heart rate monitors the Empatica E4 and the Polar H10 provide a reliable alternative for the REFA machine and can be used in research. To test the monitors, we will induce a physical and mental stress reaction to trigger an increase in heart rate and compare the feasibility and validity of the devices. The devices are tested in three conditions: a resting condition (sitting), a physical stress condition (standing) and a mental stress condition (Stroop). While resting the participants are asked to sit for one minute; When standing, we ask the participants to perform a physical task that involves getting up and standing for one minute. Finally, the participants are asked to perform a test battery consisting of two Stroop tests (one regular Stroop test and one modified emotional Stroop test).

Hypotheses

We expect the polar H10 and EE4 to detect a significant difference, i.e. a change in physiology in HR and HRV during the different conditions. We expect to see this in the in the length of the IBIs, with longer IBIs indicating a faster HR. In addition, we expect a difference in HR and HRV when standing compared to when. We also expect a significant difference in the emotional Stroop task to the regular Stroop test because the emotionally valenced words may elicit a higher emotional reaction. Finally, we expect the EE4 and the polar H10 to compare to the golden standard REFA during all tasks.

Methods

Ethical Statement

The Ethical Committee Psychology (ECP) granted ethical approval for this study, affiliated with the University of Groningen (UG), the Netherlands.

Participants

The total sample [*N*=44: women (*n*= 31), men (*n*=13)]. This study is a within-subject design and therefore has no separate control group. The participants were recruited through convenience sampling psychology students from the UG. The students were rewarded with two Sona credits, a reward system created by the UG. Students were likely to be from different nationalities, i.e. non-native speakers; however, the experiment was conducted in English. The sample was not randomized since there was only one experimental group. The participants filled out a prescreening indicating gender and whether they use optical help and handedness.

Materials

Stroop tests

A Stroop test was created. The test has two subtypes a regular Stroop test and an emotional Stroop test. Both of the tests consist of three subtests- The three subtests of the regular Stroop test are 1) a congruent Stroop, i.e. the colour of the word is congruent with the colour-word itself; the word red is presented in the colour red, 2) an incongruent Stroop test, i.e. the colour of the word is incongruent with the colour-word; the word red is presented in a different colour such as green 3) a mixed test, i.e. the trial consists of congruent and non-congruent words. The different colours presented were green, yellow, red or blue. The order of the word-colour combination was randomized in each subtest.

The emotional Stroop test is a variation of the regular Stroop test. The words are in different colours; however, emotionally valenced words are used instead. The words are either positively or negatively valenced. The negative words are associated with war, such as *kill*. In contrast, the positive words have been chosen to complement the negative word. Therefore, the counterpart of *kill* is *cure*. The words were equally distributed (*N*=32; *n-positive*= 16, *n-negative*=16). The negative words are, on average, shorter (*M*= 5.25 letters; *Mdn*= 5 letters) compared to the positive words (*M*= 5.88 letters; *Mdn*= 4 letters). A complete list of the words can be found in Appendix A. The emotional Stroop consisted of three individual subtests: 1) a positive emotional Stroop, 2) a negative Stroop test, and 3) a mixed test, i.e. both negative and positive words are presented.

Participants get positive or negative feedback to reinforce correct and punish false responses during the trial. In addition, the test has a staircase feature implemented, meaning that correct answers will be followed quicker by a new word. In contrast, wrong or no answers will lead to a slower response to the next item.

Psychomotor test

A psychomotor task was created to interrupt the primary Stroop task to assess the participants' attention. This task was presented on a second screen, placed some distance away from the Stroop task. The screen said *Alarm on standby…* . Noise appears in random intervals from approximately three to five minutes. The Alarm could be stopped by pressing the space key on the keyboard. The participants got feedback on how much time they needed from this task. Although part of the study, this paper will not assess the psychomotor test.

Apparatus

Empatica E4

The EE4 measures HR via four different sensors to measure HRV: 1) an electrode to measure Electrodermal activity, referring to the electrical conductance of skin responding to sweat. 2) 3-axis accelerometer, monitoring the heartbeat (HB). 3) a temperature sensor, and 4) a photoplethysmography, which measures blood volume pulse (Schuurmans et al., 2020). However, only photoplethysmography is an indicator of HR. The EE4 does not record HR and HRV; however, it records IBIs, which can be used to estimate HR. The EE4 has an in-built algorithm that removes incorrect peaks such as motion artefacts from the data. Incorrect peaks may be the result of intense movement (support.empatica.com, 2020). In this study, however, we will only assess IBIs as an indicator of physiological change.

Polar H10

The polar H10 is a chest strap that contains measurement sensors and transmitter units to measure HR and IBI. The polar H10 is equipped with a built-in memory system to allow to access data online. The polar submits data via Bluetooth connection. Furthermore, the raw data can be accessed through the software provided by polar (polar.com, n.d).

EEG amplifier REFA

The REFA is a multifunctional device measuring electrophysiological input such as EEG, ECG or other physiological measures. The REFA is used for medical and psychological research. The REFA includes a program that cleans the data from interferences and movement artefacts. However, raw data can always be accessed (tmsi.com/products/refa/n.d.). The REFA is a golden standard measurement device for ECG and is used as the norm in this study.

Button Box

A button box was created in-house at the UG. The box contains four different buttons in a row. The buttons include LED lights in red, blue, green and yellow (left to right).

Data Recording

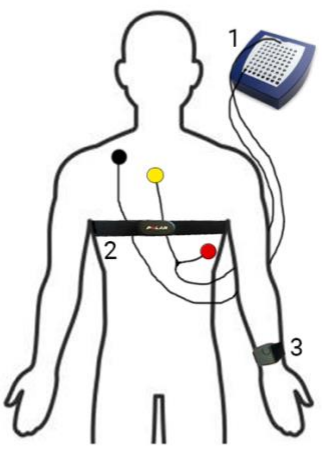
The recordings were conducted in a one-time session only. The participants wore the three devices at all times during the session. The session lasted on average 20 min.. Baseline HRV measures were obtained while the participants performed all conditions. The data was recorded using LabRecorde.exe. The data was later analyzed using MATLAB.

Procedure

Participants filled in the pre-screening and, after that, were invited to participate in the laboratory of the UG. At arrival, the participants were informed about the study rationale; this was either provided orally or in an informal letter (Appendix B). Informed consent was obtained (Appendix C), and participants were assured that their data and ECG would remain anonymous and not shared. Before starting the experiments, the participants were connected to the three devices. The EE4 is a wristband placed on either one of the hands. The polar H10 band was placed around the chest and adjusted correctly. Lastly, the electrodes were applied to the participant's chest wall. Two electrodes were placed below the collarbone on the right sight of the body. A third was placed approximately at the end of the ribcage of the left body side. A visualization of this can be found in Figure 1.

Figure 1.

Placement of Devices on Participants

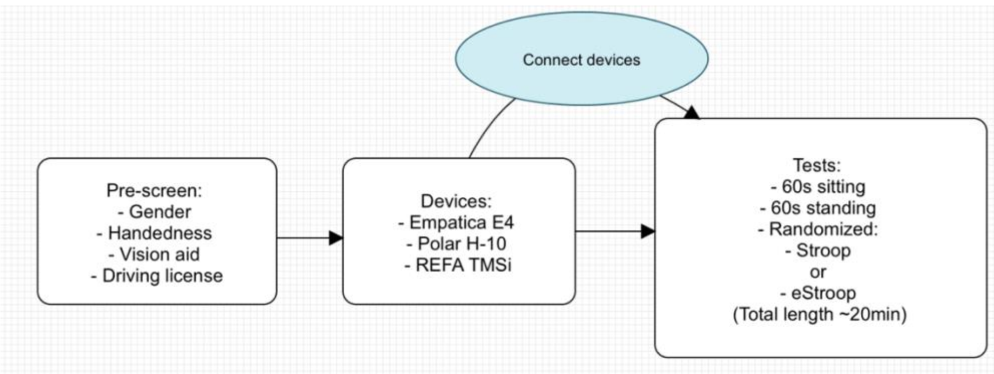


Note. 1= REFA with attached electrodes, 2= Polar H10, 3= Empatica E4.

After the devices were connected to the lab recorder, participants were instructed with the task in a separate noise-isolated room. The participants were instructed to do a physical task first. The participant started with a physical task of sitting for one minute, followed by a one-minute standing period. Afterwards, the Stroop test began. The participants are instructed to ignore the context and focus on the colour only.The order of the two Stroop types was randomized, meaning that the test started either with a regular Stroop or an emotional Stroop task. The order of the three subtypes was randomized. In addition, the participants were informed that noise would appear from a second computer in the room at any given time and that they had to stop the noise as quickly as possible. They could do so by clicking the space bar. After completing the task, the participants could leave the room, and the ECG devices were removed. Ultimately, participants were granted their credits. A timeline can be found in Figure 2.

Figure 2

Procedure of study



Data Analysis

The Analysis was performed in JASP 0.16.2 and Rstudio 2022.02.3 Build 492. To assess the devices, descriptive statistics and correlations were obtained. Descriptive statistics were assessed for the total sample and each participant individually. Additionally, a Welch Sample t-test was conducted to validate the correlation findings. For the EE4, there was no Pearson correlation and no Welch Sample t-test conducted because the data sets are not the same size and therefore cannot be compared. The Group means of the devices were compared using descriptive statistics using the REFA as the norm. To be considered the exact measurement, the recorded IBIs can deviate 0.002s from the REFA. A one-way ANOVA was performed to assess whether there was a difference between the conditions. An additional non-parametric test, the Kruskal-Wallis test, was performed when ANOVA suggested significance. The alpha level of .05 was applied in all cases.

Results

A total sample of 43 participants was assessed. Participants were excluded if the recorded data could not be used. For the analysis, 16 participants were excluded. Reasons were because the electrodes came loose during the experiment, there were too many movement artefacts, or the data analysis could not match the IBIs of the three devices. After exclusion the sample size was: [*N*=28: women (*n*= 20), men (*n*=8)]. The demographics of the sample can be found in Table 1.

Table 1

Characteristics of participants after exclusion

	Gender		**Handedness**		**Vision aid**		
	Female	*Male*	*Right*	*Left*	*No aid*	*Contacts*	*Glasses*
Total	20	8	26	2	21	5	2
Percentage (%)	71.43	28.57	92.86	7.14	75.00	17.86	7.14

Note. Participants were excluded if the recorded data could not be used

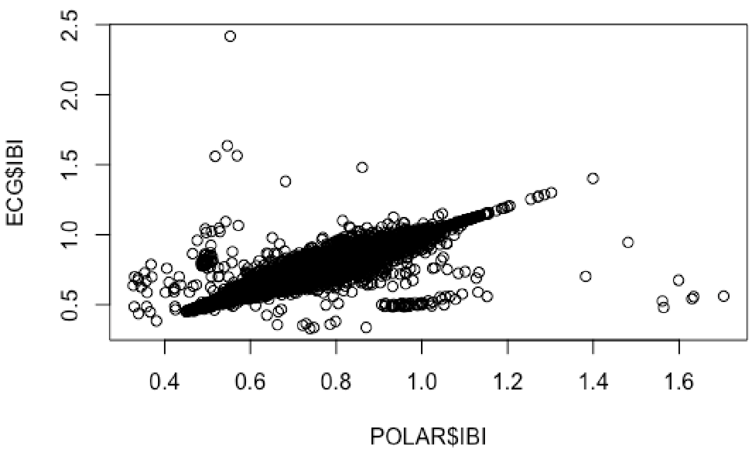
Comparing Polar H10 and Empatica E4 with the REFA

Polar compared to the REFA

To test whether the polar can be compared to the REFA, a Pearson correlation and a Welch sample t-test of the whole sample were conducted. The REFA and Polar were found to be strongly positively correlated, *r*(26) =.94, *p*<0.001, 95% CI [.941, .943]. A Q-R plot can be found in Figure 1. An independent t-test was conducted to compare the Polar and the REFA. There was no significant result found between the Polar (*M=.700, SD=.107*) and the REFA (*M=.700, SD=.107*); t(89706)=0.003, p=.997, 95% CI [-.001, .001]. The descriptive for the total sample can be statistics can be obtained from Table 2. To test whether the polar record was reliable across all subjects, the statistical descriptive were also assessed individually. The description of the individual can be obtained from Table 3.

Figure 3

Q-R plot of correlation of the REFA and POLAR



Note. ECG$IBI= IBI points measured by the REFA, POLAR$IBI= IBI points measures by the polar H10

Table 2

Statistical Descriptive for collected IBIs (Total sample)

	IBI		
	REFA (s)	**Polar H10 (s)**	**Empatica (s)**
N	44854	44854	7758
Mean	0.700	0.700	0.728
SD	0.107	0.107	0.135
IBIs (%)	100.00	100.00	17.30
Minimum	0.330	0.327	0.406
Maximum	2.419	1.704	1.844

Note.N=collected IBIs, SD= Standard deviation, IBIs (%) = percentage of IBIs measured when REFA considered to be 100%

Table 3

Statistical Descriptive for IBIs measured by participant

	Device	N	Mean	Std. Dev.	IBIs (%)	Minimum	Maximum
P001	REFA	1250	0.903	0.095	100.00	0.653	1.401
	PolarH10	1250	0.903	0.095	100.00	0.653	1.399
	EE4	345	0.940	0.116	27.60	0.656	1.391
P004	REFA	1308	0.819	0.105	100.00	0.553	1.078
	PolarH10	1308	0.819	0.105	100.00	0.553	1.078
	EE4	0	-	-	0.00	-	-
P005	REFA	1665	0.669	0.054	100.00	0.330	0.911
	PolarH10	1665	0.671	0.057	100.00	0.437	1.599
	EE4	122	0.685	0.078	7.33	0.500	0.922
P006	REFA	1566	0.661	0.059	100.00	0.518	0.961
	PolarH10	1566	0.661	0.059	100.00	0.518	0.960
	EE4	283	0.697	0.094	18.07	0.516	0.984
P008	REFA	1579	0.736	0.068	100.00	0.336	1.108
	PolarH10	1579	0.736	0.080	100.00	0.330	1.704
	EE4	21	0.766	0.079	1.33	0.641	0.953
P014	REFA	1642	0.654	0.056	100.00	0.487	0.896
	PolarH10	1642	0.654	0.056	100.00	0.486	0.895
	EE4	393	0.667	0.077	23.93	0.500	1.016
P016	REFA	1438	0.840	0.102	100.00	0.613	1.154
	PolarH10	1438	0.840	0.102	100.00	0.614	1.156
	EE4	205	0.904	0.127	14.26	0.547	1.250
P017	REFA	1525	0.720	0.056	100.00	0.571	1.206
	PolarH10	1525	0.720	0.056	100.00	0.572	1.205
	EE4	215	0.748	0.104	13.44	0.531	1.391
P018	REFA	1597	0.708	0.080	100.00	0.554	1.189
	PolarH10	1597	0.708	0.080	100.00	0.554	1.192
	EE4	292	0.747	0.095	18.72	0.547	0.984
P019	REFA	1494	0.711	0.089	100.00	0.488	1.149
	PolarH10	1494	0.711	0.089	100.00	0.489	1.151
	EE4	44	0.722	0.110	2.26	0.484	0.969
P020	REFA	1601	0.624	0.038	100.00	0.523	0.824
	PolarH10	1601	0.624	0.038	100.00	0.523	0.825
	EE4	144	0.634	0.049	8.99	0.516	0.828
P021	REFA	1657	0.666	0.052	100.00	0.535	0.899
	PolarH10	1657	0.666	0.052	100.00	0.535	0.899
	EE4	246	0.685	0.077	17.62	0.500	1.156
P025	REFA	1489	0.762	0.061	100.00	0.593	1.044
	PolarH10	1489	0.762	0.061	100.00	0.593	1.045
	EE4	423	0.783	0.086	28.31	0.563	1.078
P026	REFA	1523	0.750	0.084	100.00	0.531	1.151
	PolarH10	1522	0.750	0.084	99.93	0.533	1.151
	EE4	101	0.816	0.160	6.64	0.578	1.844
P027	REFA	1886	0.573	0.076	100.00	0.450	2.419
	PolarH10	1886	0.572	0.063	100.00	0.449	1.634

	EE4	457	0.570	0.061	24.23	0.406	0.875
P028	REFA	1790	0.641	0.047	100.00	0.511	0.858
	PolarH10	1790	0.641	0.047	100.00	0.512	0.859
	EE4	758	0.648	0.061	42.35	0.500	0.938
P029	REFA	1957	0.570	0.040	100.00	0.459	1.024
	PolarH10	1957	0.569	0.041	100.00	0.358	1.026
	EE4	395	0.575	0.044	14.92	0.438	0.734
P032	REFA	1588	0.708	0.072	100.00	0.501	0.964
	PolarH10	1588	0.708	0.072	100.00	0.501	0.964
	EE4	219	0.729	0.097	13.79	0.500	1.094
P034	REFA	1668	0.628	0.049	100.00	0.496	0.978
	PolarH10	1668	0.628	0.049	100.00	0.495	0.979
	EE4	213	0.644	0.070	12.77	0.453	0.797
P035	REFA	1611	0.711	0.080	100.00	0.511	1.481
	PolarH10	1611	0.710	0.080	100.06	0.327	1.481
	EE4	230	0.706	0.091	1.43	0.500	1.000
P036	REFA	1651	0.691	0.039	100.00	0.535	0.851
	PolarH10	1651	0.691	0.039	100.00	0.536	0.852
	EE4	100	0.692	0.051	6.06	0.578	0.797
P037	REFA	1518	0.861	0.068	100.00	0.337	1.071
	PolarH10	1518	0.861	0.068	100.00	0.329	1.070
	EE4	1081	0.871	0.062	71.21	0.656	1.203
P038	REFA	2002	0.610	0.050	100.00	0.498	0.932
	PolarH10	2002	0.610	0.050	100.00	0.499	0.932
	EE4	223	0.654	0.097	11.14	0.469	1.266
P039	REFA	1255	0.850	0.095	100.00	0.486	1.150
	PolarH10	1255	0.850	0.095	100.00	0.485	1.152
	EE4	92	0.857	0.111	7.33	0.656	1.078
P040	REFA	1505	0.720	0.050	100.00	0.425	0.868
	PolarH10	1505	0.720	0.050	100.00	0.423	0.868
	EE4	243	0.761	0.146	16.15	0.516	1.391
P041	REFA	1895	0.610	0.043	100.00	0.464	0.837
	PolarH10	1895	0.610	0.046	100.00	0.359	1.131
	EE4	391	0.628	0.072	20.63	0.438	0.875
P042	REFA	1491	0.748	0.068	100.00	0.580	1.126
	PolarH10	1491	0.748	0.068	100.00	0.579	1.127
	EE4	338	0.759	0.077	22.67	0.578	1.000
P043	REFA	1703	0.696	0.062	100.00	0.515	1.088
	PolarH10	1704	0.695	0.063	100.00	0.353	1.087
	EE4	184	0.738	0.121	10.80	1.516	1.500

Note. For P004 the EE4 did not measure any data. IBIs = IBIs measured when REFA is 100%

Empatica compared to REFA

Descriptive statistics were compared to assess the difference between the EE4 and the REFA. The EE4 did not report the same measurements

[*M*=0.728, *SD*= 0.135, *Minimum=0.406,* Maximum= 1.844] as the REFA (*M*=

0.700, *SD*= 0.107, *Minimum=0.330,* Maximum= 2.419). The EE4 recorded 17.30% of the IBIs that the REFA recorded. The descriptive for the total sample can be statistics can be obtained from Table 2. To test whether the EE4 record was reliable across all subjects, the statistical descriptive were also assessed individually for each subject. The descriptives can be obtained from Table 3.

Emotional and normal Stroop task.

As we are not interested in the individual conditions, only the Emotional Stroop and Stroop tests were assessed, and the subtest was not assessed individually. A one-way analysis of variances was performed for each of the three devices. There was no significant difference found between the emotional Stroop and the normal Stroop by the REFA ($F(1,291) =0.768, p=.382$), by the polar ($F(1,291) =0.903, p=.343$) or by the EE4 ($F(1,210) =0.266, p=.607$).The REFA recorded, on average, a slightly lower mean of IBIS in the Emotional Stroop condition (*M*=.709) than in the Stroop condition (*M*=.718). The polar recorded, on average, a slightly lower mean of IBIS in the Emotional Stroop condition (*M*=.708) than in the Stroop condition (*M*=.718). The EE4 recorded, on average, a slightly lower mean of IBIS in the Emotional Stroop condition (*M*=.732) than in the Stroop condition (*M*=.741).

Standing and Sitting

A one-way ANOVA was performed for each of the three devices to assess whether there is a difference between the conditions of Standing and Sitting. There was a significant difference found between the Standing and the Sitting condition by the REFA ($F(1,73) =20.91, P=<.001$), by the polar ($F(1,74) =21.13, P=<.001$) and by the EE4 ($F(1,66) =21.95, P<.001$). Additionally, a Kruskal-Wallis t-test suggested there was significance between the conditions by the REFA $H(1)= 195.14$, $p= <.001$, by the polar $H(1)= 195.14$, $P= <.001$ and by the EE4 H(1)= 18.43.14, $P= <.001$. The REFA recorded, on average, a higher mean in the Standing condition (*M*=.751) than in the Sitting condition (*M*=.646). The Polar recorded, on average, a higher mean in the Standing condition (*M*=.759) than in the Sitting condition (*M*=.646). Lastly, the EE4 recorded, on average, a higher IBI mean in the Standing condition (*M*=.761) than in the Sitting condition (*M*=.640). The REFA recorded more IBIS in the standing condition (*n*= 2610,*M*= .639, *SD*=.115)) than in the Sitting condition ((*n*= 2270,*M*= .739, *SD*=.136). The polar recorded more IBIS in the standing condition (*n*= 2610,*M*= .639, *SD*=.119) than in the Sitting condition (*n*= 2276,*M*= .739, *SD*=.137). The EE4 recorded more IBIs in the standing condition (*n*= 1147,*M*= .749, *SD*=.132)) than in the sitting condition (*n*= 794,*M*= .739, *SD*=.132). The EE4 recorded 50.53% of IBIs (when REFA is used as the norm) in sitting conditions and 30.42% of IBIs in the standing condition. The descriptives can be obtained from Table 4.

Table 4

Descriptive Statistic for IBIs in the Standing and Sitting Condition

	Standing			**Sitting**		
	REFA	**Polar**	**EE4**	**REFA**	**Polar**	**EE4**
N	2610	2610	794	2270	2276	1147
M	.639	0.639	.656	0.738	.739	.749
SD	.115	.115	.119	0.136	.137	.132
IBIs (%)	100.00	100.00	30.42	100.00	100.30	50.53
Min	.330	.358	.438	0.492	.327	.500
Max	2.419	1.382	1.094	1.401	1.599	1.391

Note. N= sample, M= Mean, SD= Standard Deviation, IBIs = IBIS measured in percent when REFA is 100%, Min=Minimun, Max=Maximum

Discussion

In this study, we aimed to assess whether wireless ECG devices can compare with the clinical golden standard of the ECG. In addition, we also aim to assess whether the devices can detect change during different conditions.

Polar compared to the REFA

We hypothesized that the polar is a reliable ECG and measures as accurately as the REFA, which is considered the golden standard. Per our expectations, the polar showed excellent agreeance with the REFA. A numeric comparison shows that both devices measure the same IBIs for the whole sample and individually across participants. The correlation between the polar and the REFA is robust, indicating that the polar measured the same IBIs as the REFA for most of the sample. There were minor deviations in the maximum and minimum measurements; this is most likely due to outliers in the sample causing my movement artefacts. Furthermore, the confidence interval indicated a small difference in overall means of the REFA and the polar.The findings are in line with previous research finding on High agreement between the polar H10 and the golden standard ECG machines (Martinez Ruiz, 2022). Overall the polar is highly agreeable with the REFA and can be used as alternative.

Empatica

We hypothesized that the EE4 has high agreeance with the golden standard REFA. Our findings did not support the hypothesis. The numerical comparison showed that the EE4 records, on average, longer IBIs than the REFA. The EE4 recorded, on average, much fewer IBIS than the REFA, suggesting that it missed a high number of the IBIS. This is likely due to sensitivity to movement. In the sitting condition, the EE4 recorded more IBIs than in the standing condition, suggesting that it collects more data when there is less movement. However, even in the sitting, i.e. resting condition, the EE4 recorded substantially fewer IBIs than the REFA. Even though the EE4 did not record nearly enough IBIs to be accurately compared with the REFA, the EE4 was in agreeance with the REFA in assessing the individual conditions. In agreeance with the REFA, the EE4 found a significant effect in the standing versus sitting condition, suggesting that the EE4 is able to detect change in the HRV; however, it does not measure enough IBIs to be statistically compared with the REFA.

Schuurmans et al. (2020) compared the EE4 with the Ambulatory Monitoring System (VU-AMS), i.e. a golden standard ECG machine. Even though data loss was experienced, previous studies ultimately found that the EE4 was highly comparable with the golden standard ECG when the raw data was analyzed and corrected (Schuurmans et al., 2020 & Milstein et al., 2020). This is partially in line with our findings. The comparison between the different conditions suggests that the EE4 is measuring in agreeance with the REFA; however, the EE4 misses more than 80% of the IBIs and can therefore not be accurately compared.

Impact of physical task on HR

We hypothesized that there would be a difference between the physical condition (standing) and the resting condition (sitting). In line with our hypothesis, we found a significant difference between the physical and resting conditions. Suggesting that the participants had, on average, longer IBIs in the physical condition than while resting suggesting a faster HR while standing. A significant difference was found by all three devices individually, suggesting their accuracy in measuring differences in HRV. In the present study, the physical task was a change of position. Khare and Chawala (2016) investigated the effects of body change on QRS means. They found that the majority of participants showed changes in the QRS complex. This is in line with our findings that suggest a difference in the posture standing compared to sitting.

Impact of emotional task on HR

We hypothesized that there would be a difference in HR in the emotional Stroop task compared to the regular Stroop task. We did not find evidence to support this claim. There was no significant difference between the two subtests, suggesting that participants did not have a different HR or HRV in the emotional condition compared to the normal condition. This suggests that the emotional manipulation did not affect the HR. This was found by all three devices individually, suggesting their accuracy in detecting differences in this condition. In line with this, Fackreall et al. (2013) assessed the emotional effect of dilution effect on the Stroop, hypothesizing that negative words would claim participant's attention for longer did not find a significant effect (Fackrell et al., 2013).

Strengths and Limitations

The study had several strengths Furthermore, the study was an experiment conducted in a laboratory; therefore, confounding factors that could influence participants were limited. This ensures that were measure only the effect that we intended to measure. The study design was a within-subject design. Therefore, the participants were their control group, limiting confounding factors such as individual differences in HRV between groups. While the study had these advantages, the limitation needs to be considered. Even though the study had a large sample size during the analysis almost 30% of the participants needed to be excluded resulting in only a moderate sample size. The sample was a convenience sample that participated in the study to pass a course in their study program; this limits generalizability. At the beginning of the data assessment, electrodes named *huggables* were used. These electrodes are easy to remove; however, many electrodes came loose during the experiment. This resulted in a loss of data. A limitation was also that the EE4 did report a lot less data than expected. This led to a lack of statistical analysis. Lastly, the emotional manipulation may not have been strong enough to observe an effect on HRV. Merely looking at emotional words may not elicit strong enough emotion in all participants.

Practical Implications

Due to the movement restrictions the ambulatory ECG carries, a more practicable ECG is needed. We found that the polar is highly agreeable with the golden standard devices. This emphasized that the polar can be used in clinical and research settings instead of the

REFA or other ECG machines. Using a less invasive ECG device is highly relevant for researchers and clinicians that have to deal with a restriction such as movement restrictions or the comfortability of being attached to a machine. A wireless device will likely make these restrictions easier.

Future Directions

Future studies should assess the validity of the EE4 when used on the non-dominant hand. Even though we asked for hand-dominance in the pre-screen, participants could choose where they would like to wear the watch. Future studies should also investigate the accuracy and reliability of the EE4 when corrected. Milstein et al. (2020) looked at the raw data the IBI collected and were able to extract the IBIs from there (Milstein et al. 2020). A study with similar rationales could therefore conclude the accuracy of the EE4 in different conditions. Future studies could also assess the polar H10 and the EE4 in longer duration tasks and across tasks such as more physical activity and other psychological research paradigms. Latsly, future research should investigate the effect on HR of a stronger emotional manipulation such as evaluating HR while participants are involved in emotional talks or situations.

Conclusion

The polar is highly agreeable with the clinical standard REFA and can be used in research or clinical settings where patients or participants need to move. The EE4 was not agreeable with the REFA; however, it detected significant differences between different conditions suggesting that it can detect variation in HR. Unfortunately, no conclusion can be made about the EE4 because it did not collect enough measurements. Finally, we did not find an effect on emotional manipulation of the Stroop test. Lastly, the devices found a significant effect in the standing condition compared to the resting condition suggesting that they are capable of detecting variability in HR.

References

Chattopadhyay, S., & Das, R. (2021). Comparing heart rate variability with polar H10 sensor and pulse rate variability with LYFAS: A novel study. *Journal of Biomedical Engineering and Technology*, *9*(1), 1–9.

Chua, E. C.-P., Tan, W.-Q., Yeo, S.-C., Lau, P., Lee, I., Mien, I. H., Puvanendran, K., & Gooley, J. J. (2012). Heart Rate Variability Can Be Used to Estimate Sleepiness-related Decrements in Psychomotor Vigilance during Total Sleep Deprivation. *Sleep*, *35*(3), 325–334. https://doi.org/10.5665/sleep.1688

de Luna, Antoni., Fiol-Sala, Miquel., Bayés-ís, Antoni., & Baranchuk, Adrian. (2022). *Clinical Electrocardiography A Textbook.* (5th ed., Vol. 1–1 online resource (611 p.)). John Wiley & Sons, Incorporated; WorldCat.org. http://public.eblib.com/choice/PublicFullRecord.aspx?p=6725094

de Luna, A. B., Batchvarov, V. N., & Malik, M. (n.d.). *The Morphology of 1 the Electrocardiogram*. 36.

Gilgen-Ammann, R., Schweizer, T., & Wyss, T. (2019). RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise. *European journal of applied physiology*, *119*(7), 1525–1532. https://doi.org/10.1007/s00421-019-04142-5

Fackrell, K., Edmondson-Jones, M., & Hall, D. A. (2013). A controlled approach to the emotional dilution of the Stroop effect. *PloS one*, *8*(11), e80141. https://doi.org/10.1371/journal.pone.0080141

Khare, S., & Chawala, A. (2016). Effect of change in body position on resting electrocardiogram in young healthy adults. *Nigerian Journal of Cardiology*, *13*(2), 125. https://doi.org/10.4103/0189-7969.187711

Karthikeyan, P., Murugappan, M., & Yaacob, S. (2014). Analysis of Stroop Color Word Test-

Based Human Stress Detection using Electrocardiography and Heart Rate Variability Signals. *Arabian Journal for Science and Engineering*, *39*(3), 1835–1847. https://doi.org/10.1007/s13369-013-0786-8

Kim, H.-G., Cheon, E.-J., Bai, D.-S., Lee, Y. H., & Koo, B.-H. (2018). Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature. *Psychiatry Investigation*, *15*(3), 235–245. PubMed. https://doi.org/10.30773/pi.2017.08.17

Kunkels, Y. K., Roon, A. M., Wichers, M., & Riese, H. (2021). Cross-instrument feasibility, validity, and reproducibility of wireless heart rate monitors: Novel opportunities for extended daily life monitoring. *Psychophysiology*, *58*(10). https://doi.org/10.1111/psyp.13898

Luque-Casado, A., Zabala, M., Morales, E., Mateo-March, M., & Sanabria, D. (2013). CognitivePerformance and Heart Rate Variability: The Influence of Fitness Level. *PLOS ONE*, *8*(2), e56935.

Martinez Ruiz, Armando, "Validity and Reliability of the Apple Series 6 and 7 Smartwatches and Polar H-10 Monitor on Heart Rate" (2022). *Open Access Theses & Dissertations*. 3517.https://scholarworks.utep.edu/open_etd/3517

Milstein, N., & Gordon, I. (2020). Validating Measures of Electrodermal Activity and Heart Rate Variability Derived From the Empatica E4 Utilized in Research Settings That Involve Interactive Dyadic States. *Frontiers in behavioral neuroscience*, *14*, 148. https://doi.org/10.3389/fnbeh.2020.00148

Pinel, J. P. J. (2020). *Biopsychology, 11th ed.*. Harlow, Essex: Pearson Education Limited.

Prokhorov, M. D., Karavaev, A. S., Ishbulatov, Y. M., Ponomarenko, V. I., Kiselev, A. R., & Kurths, J. (2021). Interbeat interval variability versus frequency modulation of heart rate. *Physical Review E*, *103*(4), 042404. https://doi.org/10.1103/PhysRevE.103.042404

Satish, P., Muralikrishnan, K., & Balasubramanian, K. (n.d.). Heart rate variability changes during stroop color and word test among genders. *Indian J Physiol Pharmacol*, 8.

Schuurmans, A. A. T., de Looff, P., Nijhof, K. S., Rosada, C., Scholte, R. H. J., Popma, A., & Otten, R. (2020). Validity of the Empatica E4 Wristband to Measure Heart Rate Variability (HRV) Parameters: A Comparison to Electrocardiography (ECG). *Journal of Medical Systems*, *44*(11), 190. https://doi.org/10.1007/s10916-020-01648-w

Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health*, *5*, 258. https://doi.org/10.3389/fpubh.2017.00258

Stamou, S. S., Mamali, A., & Papageorgiou, E. (2020). Impact of Physical Activity on Heart Rate, Blood Pressure and Rate-Pressure Product in Healthy Elderly. *Health Science Journal*, *14*(2), 712.

Taelman, J., Vandeput, S., Spaepen, A., & Van Huffel, S. (2009). Influence of Mental Stress on Heart Rate and Heart Rate Variability. In J. Vander Sloten, P. Verdonck, M. Nyssen, & J. Haueisen (Eds.), *4th European Conference of the International Federation for Medical and Biological Engineering* (pp. 1366–1369). Springer Berlin Heidelberg.

Vazan, R., Filcikova, D., & Mravec, B. (2017). Effect of the Stroop test performed in supine position on the heart rate variability in both genders. *Autonomic Neuroscience*, *208*, 156–160. https://doi.org/10.1016/j.autneu.2017.10.009

Websites

Polar.com (retrieved 01. April. 2022) Polar H10 Heart Rate Sesnor, https://www.polar.com/us-en/products/accessories/h10_heart_rate_sensor

Empatica.store (retrieved 01. April. 2022) E4 Wristband, https://store.empatica.com/products/e4-wristband?variant=17719039950953

Empatica store (retrieved 27.May.2022), E4 data-IBI Expected signal https://support.empatica.com/hc/en-us/articles/360030058011-E4-data-IBI-expected-signal

TMSi (retrieved 26. May. 2022), REFA amplifier, https://www.tmsi.com/products/refa/

Appendix A: Word List for Emotional Stroop Task

Negative words	Positive words
War	Peace
Attack	Hug
Pain	Happy
Blood	Healthy
Bomb	Flower
Kill	Cure
Gun	Toy
Death	Life
Anger	Friendly
Nuclear	Sunshine
Panic	Calmness
Anxiety	Excited
Combat	Ally
Enemy	Friends
Explosion	Celebration
Fighting	Love

Appendix B: Information Sheet

university of groningen / faculty of behavioural and social sciences

INFORMATION ABOUT THE RESEARCH

VERSION FOR PARTICIPANTS

"MEASURING STRESS: CROSS-INSTRUMENT VALIDITY AND REPRODUCIBILITY OF HEART RATE MEASUREMENT DEVICES?"

EC code

Why do I receive this information?

- You are invited to participate in this study because you are a first year Psychology student at the university of Groningen.
- The researchers involved in this study are
  - Mark Span, PhD, BSS/Psychology, principal investigator and supervisor: processing, analysis, retention, sharing, and publication.
  - Rover Willemars
  - Theres Patzelt
  - Nienke Buist
  - Harmien Tamsma
  - Lisanne Zondag , bachelor students: data collection, processing, analysis, retention, sharing, and publication.

Do I have to participate in this research?

- Participation in the research is voluntary. However, your consent is needed. Therefore, please read this information carefully. Ask all the questions you might have, for example because you do not understand something. Only afterwards you decide if you want to participate. If you decide not to participate, you do not need to explain why, and there will be no negative consequences for you. You have this right at all times, including after you have consented to participate in the research.

Why this research?

- Heartrate and more specifically its variability is a physiological marker for perceived stress. Whenever we are exposed to potential risks, a psychological fight-or-flight response is triggered, which makes the heartrate spike. Chronic stress is associated with a reduction of the variability of the heartrate. (e.g., Tegegne et al, 2018) Recently, a new generation of compact, and relatively inexpensive heart rate monitors have become available. Many of these devices have closed-source parameter estimation algorithms. Therefore, it is unsure what the values obtained by these devices mean in research. We want to test three devices that give access to the raw ECG data, and compare the values obtained for Interbeat Intervals (IBI) for validity and reproducibility. We want to test the Polarband H10 and the Empatica E4 wristband against the golden standard (ECG). In the process of doing so, we want to use a set of work stress inducing tasks.
- The tasks used to induce work stress are standard and emotional Stroop tasks, combined with a motor vigilance task. The tasks are used to induce involvement in the subjects in performance of the task, thereby affecting the heart rate.

➢ What do we ask of you during the research?

- Before you begin the experiment, you will be asked to sign an informed consent form.
- You will be asked to sort various words according to either their meaning, their colour of their emotional loading. Your task can be interrupted by an alarm you will have to turn off.You will have a chance to familiarize yourself with the task during a set of practice trials.The total time for the research will be one hour. The introduction, familiarization with the study and consent will take about 15 minutes, and the experiment will last about 45 minutes.
- You will be compensated for your time with 2 SONA credits.

➢ What are the consequences of participation?

- By participating in this study, you will contribute the development of the field of cognitive neuroscience by broadening the understanding of various heartrate measurements are correlated.
- There are no anticipated disadvantages or risks related to participating in this study.

➢ How will we treat your data?

- Data processing takes place exclusively for educational purposes (Bachelor thesis).
- The data that will be processed will be the heartrate data and the behavioural data from the experiment (your keypress responses). No sensitive personal data will be processed. There are no data processing risks anticipated.
- For the purposes of granting credits, you are identified by your SONA ID. However, the experimental data will not contain your SONA ID, you will be given a number (for data storage purposes) which is not linked to your personal information. The research data will be securely stored in the university Y drive.
- Your data will be stored for 10 years (in accordance with the Faculty Data Storage Protocol).
- The experimental data may be later shared upon request for follow-up studies. This data will not include any personal information.

➢ What else do you need to know?

You may always ask questions about the research: now, during the research, and after the end of the research. You can do so by speaking with one of the researchers present right now or by emailing (m.m.span@rug.nl) or phoning (+31 50 36 36402) one of the researchers involved.

Do you have questions/concerns about your rights as a research participant or about the conduct of the research? You may also contact the Ethics Committee of the Faculty of Behavioural and Social Sciences of the University of Groningen: ec-bss@rug.nl.

Do you have questions or concerns regarding the handling of your personal data? You may also contact the University of Groningen Data Protection Officer: privacy@rug.nl.

As a research participant, you have the right to a copy of this research information.

Appendix C: Informed Consent Form

university of groningen

faculty of behavioural and social sciences

INFORMED CONSENT

MEASURING STRESS: CROSS-INSTRUMENT VALIDITY AND REPRODUCIBILITY OF HEART RATE MEASUREMENT DEVICES

EC code

- I have read the information about the research. I have had enough opportunity to ask questions about it.

- I understand what the research is about, what is being asked of me, which consequences participation can have, how my data will be handled, and what my rights as a participant are.

- I understand that participation in the research is voluntary. I myself choose to participate. I can stop participating at any moment. If I stop, I do not need to explain why. Stopping will have no negative consequences for me.

- Below I indicate what I am consenting to.

<u>Consent to participate in the research:</u>

[ ] Yes, I consent to participate; this consent is valid for a year.

[ ] No, I do not consent to participate

<u>Consent to processing my data:</u>

[ ] Yes, I consent to my data being used as mentioned in the research information sheet.

[ ] No, I do not consent to the processing of my data.

Participant's full name:	Participant's signature:	Date:

Full name of researcher present:	Researcher's signature:	Date:

The researcher declares that the participant has received extensive information about the research.

You have the right to a copy of this consent form.

Appendix D

Table 5

Characteristics of participants

	Gender		Handedness		Vision aid		
	Female	*Male*	*Right*	*Left*	*No aid*	*Contacts*	*Glasses*
Total	31	13	43	3	32	8	4
Percentage (%)	70..45	29.55	93.18	6.82	72.73	18.18	9.09

Note. Demographics before patients were excluded

Table 6

Mean of IBIs in the Stroop conditions

	Stroop	Emotional
EE4	.741	.732
Polar	.718	.708
REFA	.718	.709

Table 7

ANOVA comparing Emotional Stroop test with Stroop test measured by the Polar

	df	Sum of squares	Mean square	F-Value	p-level

Emotional	1	0.007	0.007	0.903	.343
Residuals	291	0.719	0.009		

Table 8

ANOVA comparing Emotional Stroop test with Stroop test measured by the REFA

	df	Sum of squares	Mean square	F.Value	p-level
Emotional	1	0.006	0.006	0.768	.382
Residuals	291	2.43	0.008		

Table 9

ANOVA comparing Emotional Stroop test with Stroop test measured by the EE4

	df	Sum of squares	Mean square	F-Value	p-level
Emotional	1	0.004	0.004	0.266	.607
Residuals	210	3.113	0.015		

Table 10

Mean of IBIs in the resting and physical conditions

	Sitting	Standing
EE4	.640	.761
Polar	.646	.750
REFA	.646	.751

Table 11

ANOVA comparing Sitting with Standing Task measured by the Polar

	df	Sum of squares	Mean square	F-Value	P-level
Sitting	1	0.205	0.205	21.13	<.001
Residuals	74	0.719	0.009		

Table 12

ANOVA comparing Sitting with Standing Task measured by the Polar

	df	Sum of squares	Mean square	Value	P-level
Sitting	1	0.204	0.205	20.91	<.001
Residuals	73	0.713	0.009		

Table 13

ANOVA comparing Sitting with Standing Task measured by the EE4

	df	**Sum of squares**	**Mean square**	**F-Value**	**P-level**
Sitting	1	0.247	0.247	21.95	<.001
Residuals	66	0.7422	0.011		