Exploring the Usage of Commercial Bio-Sensors for Multitasking Detection

Jailan Salah
German University in Cairo, AlMakinah
Cairo, Egypt
jailan.salah@gmail.com

Yomna Abdelrahman
University of Stuttgart, Bundeswehr University
Munich, Germany
yomna.abdelrahman@vis.uni-stuttgart.de

Yasmeen Abdrabou
German University in Cairo
Cairo, Egypt
yasmeen.abdrabou@guc.edu.eg

Khaled Kassem
German University in Cairo
Cairo,Egypt
khaled.kassem@guc.edu.eg

Slim Abdennadher
German University in Cairo
Cairo,Egypt
slim.abdennadher@guc.edu.eg

ABSTRACT
Most of the current adaptive systems support single task activities. The rise in the number of daily interactive devices and sources of information made multitasking an integral activity in our daily life. Affect-aware systems show exciting potential to support the user, however, they focus on the induced effect of an additional task in terms of cognitive load and stress, rather than the influence of the number of tasks i.e. multitasking. This paper presents indicators of the number of tasks being performed by the user using a set of bio-sensors. A preliminary user study was conducted with two follow-up explorations. Our findings imply that we can distinguish between the number of tasks performed based on high-end as well as cheap Heart Rate sensors. Additionally, tasks number correlates with other signals, namely wrist and forehead temperature. We provide empirical evidence showing how to differentiate between single- and dual-tasking activities.

CCS Concepts
•**Human-centered computing** → **User studies;**

Author Keywords
Multitasking, Affective Computing, Galvanic Skin Conductance, Heart Rate, Thermal Imaging

INTRODUCTION
The wide range of daily devices, tools and information sources made multitasking an integral daily activity in our life. However, our mental and cognitive performance degrades while performing more than one task simultaneously [35]. On the other hand, the continuous and increasing sources of information, as well as side tasks might be having a positive effect on the achievement of the primary task in hand.

Current ubiquitous sensors and technologies focus on supporting users while performing focused tasks [4, 19, 20]. However, single focused tasks nowadays are rarely experienced, users are often interrupted by side channels of information e.g. notifications [53, 56]. On the other hand, there is a vast increase in technological support of multitasking e.g. large displays. This drew the interest of the researchers to explore how to build interfaces that support multitasking [30]. However, in order to build systems that consider, support and adapt to the number of tasks in hand, a clear understanding of how multitasking influences our state is required.

Previous work showed that multitasking can trigger changes in the mental and cognitive needs and states of the user [11, 16]. There has been an obvious need to involve affect-aware systems in many applications [4, 47]. Different approaches include subjective analysis of user's state and assessment of physiological signals. The tendency towards using unobtrusive bio-sensors has seen a recent increase due to their feasibility of usage outside the labs. Advances in miniaturization and mass production have brought down the prices of these sensors making consumer-grade trackers readily available. Hence, investigating how to sense internal states in a robust, accurate, timely, and unobtrusive way is yet an open challenge.

Where there has been vast amount of work to estimate stress [33, 14, 41, 43], and cognitive load [36, 25, 49, 40]. We aim to estimate the internal psychological state of the user during multitasking itself rather than the induced changes, by using commercial bio-sensors. In this paper, we describe a preliminary study and two follow up explorations to assess the internal state of the user that occurs due to experiencing multitasking. Using subjective measures might not be informative about the user's state due to the timing, as measures are usually collected after the task performance. Moreover, bio-

MUM '18, November 25–28, 2018, Cairo, Egypt

DOI: https://doi.org/10.1145/3282894.3282900

signals are involuntary in nature and are hard to fake. Hence, commercial unobtrusive bio-sensors can be advantageous in detecting multitasking. Moreover, it ensures their applicability in scalable real life situations, as well as their reach to the economically-challenged groups.

We present a method for estimating the number of tasks currently performed (single vs. dual tasks) based on the bio-data of the users using commercial bio-sensors. Our method works with off-the-shelf hardware and is applicable for ubiquitous computing environments. It utilizes one of the cheapest sensors set commercially available, which opens up new opportunities for large scale deployments of state-aware technologies, as well as the involvement of the developing countries in the research field.

In our work, we advance the state-of-the-art of automatic multitasking estimation through the following contributions:

- We propose a method for differentiating between single- vs. multi-tasking by computing a set of bio-signals including, heart rate, heart rate variability, galvanic skin response and skin temperature.
- We demonstrate the validity of our metrics through a user study, showing that single- vs. multi-tasking activities strongly correlate with bio-signals.
- We validate the afford-ability in terms of costs of utilizing and deploying these sensors in educational institutes in developing countries.
- We release our dataset as open source for future researchers to build upon, replicate, and extend [1].

BACKGROUND AND RELATED WORK

Over the past years -as technology and HCI became in demand-, researches were carried out as a mean to detect emotions so that computers can learn to interact with humans, and customize themselves according to their needs. Biofeedback technology comes in all colors, starting with wearable sensors that can be used day to day, to wired equipment that is only to be used in laboratories.

Human emotions and cognitive states tend to show themselves in many cues starting from posture, gestures, speech and text [46], to the rhythm of key strokes, tone of voice [17], heart rate, heart-rate variability, skin conductivity [19], eye activity [23, 24], facial-electroencephalogram [55] and skin temperature [34, 4]. Devices used in physiological signal based state recognition systems range from contact-free cameras [32], and tracking bands, to Electromechanical film (EMFi) chairs [6]. Researchers have investigated the effectiveness of different bio-signals measurements in detecting users' states. Our work builds on three strands of prior work: (1) indicators for user's internal state to (2) measure mental workload and (3) how low-cost devices could be used to detect user's states.

Bio-Signals as Internal State Indicators

Intuitively, physiological changes can indicate the onset of certain internal states [4, 19], e.g. a dangerous situation that triggers a sense of fear causes a fight-or-flight response in which heart rate is elevated, sweat gland activity is activated which increases skin conductance, and the distribution of heat over regions such as the face change. Using the appropriate sensors to measure and record changes that occur in response to stimuli is therefore an appropriate way of automatically detecting different mental and emotional states.

Heart Rate (HR) is an important factor in affective computing. The spontaneous rhythm of the heart shows its ability to adapt to and perform in different situations [18]. HR changes from beat to beat; it increases upon inhalation, and decreases upon exhalation. Therefore, there exists no fixed heart rate for a person, but rather a naturally occurring series of irregularity in its rate, which is also known as heart rate variability [33]. Accordingly, the more the heart's rhythm fluctuates from the baseline, the better an individual is able to cope with physical and psychological strain, such as that caused by stress [5, 39]. This means the less regular the heartbeat is, the higher the heart rate variability, and vise versa.

Heart Rate Variability (HRV) is used to measure a person's Autonomic Nervous Systems (ANS) activity, which controls involuntary actions such as heartbeat rates, breathing rate, blood pressure, and sweating [49]. The Sympathetic Nervous System (SNS) is referred to as the fight and flight system, where any activity in it results in altering the heart rate. HRV measures are calculated from the respiratory rate, and classified into multiple components such as Low Frequency (LF), High Frequency (HF), Low Frequency to High Frequency (LF/HF) power ratio, and RR beat-to-beat interval [50].RR intervals are small changes (milliseconds) in the intervals between successive heartbeats, this is different from heart rate, which just averages the number of beats per minute. According to [39], LF and HF components change upon sympathetic activity. Moreover, LF/HF ratio reflects sympathovagal interaction without defining the individual contribution of each component of the ANS [31]. Finally, the RR interval reflects the time taken to complete a cardiac cycle. The higher the heart rate, the smaller the magnitude of the variation in RR interval [12, 26].

Galvanic Skin Response (GSR) is an autonomic physiological signal that is extracted from the level of sweat in the skin. It is one form of Electrodermal activity (EDA) that reflects the skin's ability to conduct electricity, and considered as a credible measurement of affect state [7, 19]. It is considered to be a useful index of changes in sympathetic arousal that are tractable to cognitive states, as it is the only autonomic psychological variable that is not contaminated by parasympathetic activity [10]. Research and empirical data have long linked GSR and its variation to changes in autonomic arousal as well as SNS changes [45].

Skin Temperature is an effective indicator for objectively evaluating human sensations, because it is controlled by sympathetic nerve activity which reflects the course of information processing in the brain. When the SNS is activated, muscles tense up and blood pressure increases, upon which blood flow is diverted from different parts of the body like fingertips, toes or the digestive tract and supplied to the vital parts, where the

[1] https://github.com/affective17/Multitasking-Detection

body will sacrifice outer limbs to feed blood to more important organs like the brain, which leads to the skin temperature decreasing or increasing according to the blood flow [40, 43]. Moreover, there is a high correlation among user's state and skin temperatures of nose and forehead [37, 4].

Informed by the literature, various bio-signals of the human body like HR, HR, GSR, and skin temperature are connected to the ANS and SNS activity. Hence, observed changes in these involuntarily changing metrics can analyzed to give insights about user's affect. In this work, we aim to utilize these indicators to investigate and explore user's state during multitasking as opposed to single task.

Measuring User's Mental State

User's viewpoints are either categorical, or dimensional [51]. They can be characterized in terms of judged valence and arousal, where valence measures whether the current state is positive or negative while arousal indicates whether the user is calm or triggered [29]. For example, stress is defined by increased arousal and negative valence. Reliably and automatically recognizing mental load can be of much help in many contexts, where arousing situations are managed and their effects can be minimized.

HR: Many studies have explored the possibility of state detection using HR measurements. Taelman et al. [50], [37] and [33] suggest that HR changes significantly with respect to the mental state. On the other hand, McDuff et al. [32] found no significant difference between HR means in the rest and cognitive load induced by performing tasks conditions. Moreover, many studies explored the possibility of detecting different states using unobtrusive wearable sensors and reported an increase in HR mean upon exposure to mental tasks [54, 37, 41, 14], while Kranjec et al. [28] found that HR decelerated in response to negative stimuli as compared with responses to positive and neutral ones.

HRV: Burns et al. [12] reported that HR and HRV are useful in measuring valence, where negative emotions have a stronger influence on them compared to positive or neutral ones. Taelman et al. [50] investigated the effect of task based mental load on HRV components. Results showed that mean RR interval decreased significantly, with a tendency for elevated (LF/HF) ratio during performing the task compared to the rest condition. Moreover, Melillo et al. [33] reported that HRV features measurements increase as an indication of low resiliency. McDuff et al. [32] addressed remote measurement of HR and HRV changes using a digital camera, where Stroop color tasks [22] were used to induce and classify restful versus stress states. Results showed that HRV was a stronger predictor of user's state compared to HR ratings, as LF and (LF/HF) components increased during performing the task. Choi et al. [14] used a time series heart monitor to explore the effect of single tasks on HRV measurements. Results indicated that RR beat to beat interval changes upon exposure to the Stroop task, with an increase in its mean. Salai et al. [41] used the same test to serve as a source of mental stress, where HF and LF features increased, while mean RR decreased.

GSR: Many studies examined GSR signals for user state measurement during performing tasks. Different studies reported increase in GSR during performing tasks that require mental load [37, 9, 36, 25, 49, 27]. Khawaji et al. [25] investigated how stress is associated with cognitive load and trust. They concluded that high mental load or low interpersonal trust can induce stress, and consequently, results in increased GSR. Additionally, Sun et al. [49] presented a multimodal approach to model states affected by both mental and physical activities using accelerometer, Electrocardiography (ECG) and GSR sensors. They reported that both HR and GSR increase with performing activities.

Author	Method	Heart Rate	HRV	GSR	Results	Conclusion
McDuff et al.	Mental Arithmetic	✓	✓	✓	90% RR 80% LF Power ratio ↑	HR alone not enough
Taelman et al.	Mensa IQ test	✓	✓		Mean RR ↓ Power Ratio ↑	
Nourbakhsh et al.	Mental Arithmetic			✓	GSR ↑	GSR highly subjective
Sun et al.	Stroop test & Mental Arithmetic		✓	✓	HF HR LF GSR ↑ LF/HF RR ↓	HR & GSR enough

Figure 1: Previous studies on mental workload recognition using bio-signals and their results

Skin Temperature: Thermal imaging has penetrated the HCI field recently [1, 2, 3]. Thermal imaging is an applicant methodology for non-contact human state appraisal. It has the capability of quantifying blood stream induced by different mental or physical activities. Puri et al. [40] explored the enthusiastic conditions of users finishing an adaptation of the Stroop color test using thermal imaging. Facial features were extracted with the forehead chosen to be the Region of Interest (ROI). Results gave confirmation that thermal imaging is a reasonable technique for measuring single task based load. Hence, studies exploring different load levels [4], with different hardware standards and cost would be beneficial for many affect-aware systems applications.

Due to the various applications of mental state assessment in affect-aware systems, a co-found has been established by extensive research for mental load that is provoked by task accomplishment. Literature in this context offered support for the applicability of defining task based mental load using bio-signals. However, most of the studies elicited mental workload using different methods, alternating in the nature of the tasks, their number, and objectives, with the aim of using such methods as a tool to elicit cognitive load or stress, then testing how the stressful state affects physiological signals. They focused mainly on monitoring and supporting the user during focused task and the additional task was used to elicit stress/cognitive load.

With the age of ubiquitous computing and multiple devices integrated in our daily life, the need to investigate the influence of the number of performed tasks is required. In this work, we are interested in analyzing and comparing the effect of different number of tasks on the user's state rather than using an additional task to elicit a stressful or cognitively demanding state. Accordingly, we are investigating this matter on an abstract level, where we start by a preliminary analysis of user's states while considering one factor, namely the number of tasks (*single vs. dual*).

State Detection using Low-Cost Sensors
Studies have investigated detecting mental state using different low-cost bio-sensors. Sun et al. [41] performed a validation study to compare a low cost HR sensor against a high standard device for stress detection. Test results approved the reliability of the affordable sensor in detecting features that change upon performing single task, namely the Stroop color test. Choi et al. [14] managed to assess mental load induced by a single task condition using a consumer-grade HR transmitter *(Polar T31)* and a custom-made respiration sensor. On the other hand, Ollander [37] investigated how well a wearable sensor performs in stress detection. Synchronized signals from a monitoring wristband, namely *(*Empatica E4) were compared against stationary laboratory equipment. Results showed that output signals do not correspond very well to each other, specially for the GSR measurements.

Contact-less sensors are advantageous to use with affect aware systems. McDuff et al. [32] showed that HRV components can be remotely captured using a low-cost digital camera and used to detect cognitive load. Additionally, Bousefsaf et al. [9] introduced a low-cost framework for detecting workload changes, where they exhibited a strong correlation between the trends of a webcam and contact skin conductance traces recognizing the task performing state.

As informed by the literature, most of the work done in low cost sensors addresses scalability of using them in mental load detection (Figure1). In this work, we are aiming at having a clear understanding of how multitasking influence user's state through investigating the effect of number of tasks being held at a time on the bio-signals of the user. Moreover, we want to collect the required bio-data using affordable devices with the aim of focusing on the feasibility of utilizing this technology in educational institutes in developing countries.

DETECTING MULTITASKING USING BIO-SIGNALS
Bio-signals is particularly promising for inferring user states for several reasons. They are readable using commercial sensors in both wearable and unobtrusive manner. In this work, our aim is to differentiate between different user's states due to number of tasks, namely zero, single and dual task conditions through bio-signals. An initial exploration was conducted using high end HR sensor.
We compared between a commercial and a high state of the art sensor results with the aim of validating the usage of commercial sensors to be useful for all learning groups, specially those in developing countries i.e economically challenged.

Additionally, a set of unobtrusive commercial bio-sensors was chosen with the aim of testing the effect of the tasks on different bio-signals so that users can use them during the day, and not just be limited to the lab usage. We decided to hold the comparison between high and low end sensors in one bio-signal only (Heart Rate) because of the feasibility of finding both sensors and buying them in our developing country. In summary, in this work we focus on the following research questions:

1. Can we distinguish between single and multitasking (dual-task) using high end commercially available bio-sensor? More specifically, do the changes in heart rate recorded by high end sensor correlate with the nature of the task ? (**RQ1**)
2. Can we still distinguish between single and multitasking when using the cheap version of the sensors? (**RQ2**)
3. Are other signals from commercial sensors -namely heart rate variability, galvanic skin response and skin temperature-capable of distinguishing between the number of tasks as well? (**RQ3**)

STUDY
To answer our research questions and to test our hypothesis of using the bio-signals to elicit multitasking, we conducted a user study in which we recorded the participant's heart rate, heart rate variability, galvanic skin response and skin (nose, forehead and wrist) temperature during three tasks:

1. Relaxing as the baseline.
2. Single-task activities.
3. Dual-task activity.

Design
We designed our study as a repeated-measures design. We studied the effect of the number of concurrent tasks on the heart rate. For the baseline we asked the participants to relax.

Single Task
The single tasks we provided five different tasks, We chose these tasks for their simplistic single nature:

1. recalling the numbers in a list in both reversed and original order,
2. recalling the *n*th number of the same list,
3. performing an arithmetic task,
4. sorting a list of names in an alphabetical order,
5. recalling the alphabet then reversing it.

Multitask
For the multitasking task, we merged both the Stroop color test [22] and arithmetic tasks together. We used a mash-up of the number test used by Orllander [37] and the additional task used in Bousefsaf's [9]. Participants were asked to subtract 13 from 1022 continuously while solving the Stroop color test. A countdown is started once they are given the color test, as well as a constant reminder to answer the mathematical question. The arithmetic tasks were provided orally by the experiment moderator, and they were asked to answer orally as well. The Stroop color test was given to them on the laptop screen. This

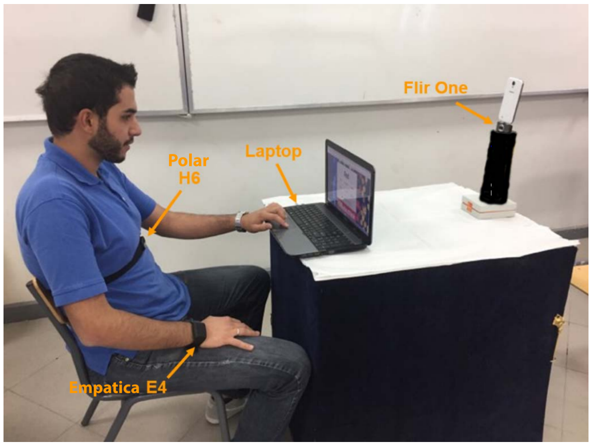


Figure 2: Study setup consisting of participant performing the tasks.

way , the participants were required to think about the correct answer for the mathematical task before it times out while solving the stroop color test which was timed as well.

To overcome the order effect of the repeated-measures experimental design, the order of the tasks was counter-balanced using a Latin Square.

Apparatus

Our experimental setup consisted of an HP Notebook -14-ac115tx laptop as shown in Figure 2. Participants were wearing the *Polar H6* chest belt [21] measuring HR, and the *Empatica E4* wristband [48] measuring HR, HRV, GSR and wrist temperature. It is worth noting that *Empatica E4* is embedded with a photoplethysmography sensor of 64 Hertz sampling frequency. It measures changes in light absorption that provides blood volume pulse from which HR and HRV can be derived. Thus, it is considered to be a high end commercial sensor that costs 1448.84 Euros, with 28 times the price of the commercial *Polar H6* sensor that costs around 61.63 Euros.

Additionally, a thermal camera (*FLIR One* android [2]) measuring facial skin temperature was attached to a phone placed one meter away from the participant. The optical resolution of our camera was 160×120 pixels. It is able to measure temperatures between -20°C and 120°C, and operates with a thermal sensitivity of 0.18°C. It is a commercial camera that costs 170.95 Euros, which is the cheapest thermal camera in the market compared to the high end thermal cameras that cost starting from 300 to 120,000 Euros.

The wavelengths captured by the camera are in the spectral range between 8μm and 14μm. The lens we use provides a 46°× 35° field of view. The thermal camera is charged and operates via male micro USB connected to an android phone, and the images are stored on the mobile and then transferred for analysis. It provides temperature information in the form of 16-bit color values encoding the temperature information. The participants were asked to look to the front facing the thermal camera placed at 1m from the participants and the screen as shown in Figure 2.

[2] http://www.flir.com/flirone/android/

Data Collection

To answer the research questions, we built a system that captures the reading from the attached sensors and a data processing software that recognizes and analyzes the user's heart rate, heart rate variability, galvanic skin response and skin temperature as described below.

Heart Rate

HR data was collected from both the *Polar H6* chest belt and the *Empatica E4* wristband. Analysis was performed on both data, with the H6 data being ran through a simple algorithm that detects inconsistencies. Inter beat interval files were inputted to Kubios HRV Standard (ver. 3.0.0) [3], and medium artifact correction was carried out, which outputted both HR and HRV data.

Galvanic Skin Response

Electrodermal activity is how an individual's sweat glands react. It is decomposed into phasic and tonic components, with the phasic component rapidly changing (known as the skin conductance response), and the tonic one slowly changing (known as the skin conductance level). *Empatica E4* produces EDA files which contain both components. Hence, analyses were performed to separate both components, and extract phasic data. Moreover, averaging across the whole signal gives inaccurate results for GSR [8]. Therefore, decomposition was performed where averages were then taken for each phase period, followed by taking means for each task.

Heart Rate Variability

HRV has many components such as RR intervals, HF, LF and LF/HF power ratio. Different analysis methods can be used to analyze HRV signals such as time-domain methods, frequency domain methods, and geometrical methods [13]. Time domain analysis is commonly used for long-time data recordings, while the frequency domain is commonly used for short-term recordings [52]. Since each task in our experiment lasted for five minutes, and geometrical methods are not significantly affected by changes in the breathing rate [38], frequency domain method was more suitable to use.

Frequency domain analysis can be decomposed into different methods, with its two most common methods being auto-regressive spectral estimation and Fourier techniques [15]. According to [42, 31], auto-regressive technique is parametric model that produces smoother spectral components, but at the same time might remove valuable data. Hence, all spectral components were obtained in each subject using the Fast Fourier Transform (FFT) method on Kubios HRV (ver. 3.0)3, where all input data was subjected to medium artifact correction. Outputs are HF power, LF power, LF/HF power ratio and RR beat to beat intervals.

Skin Temperature

Our skin temperature metrics included the wrist, nose and forehead temperatures.

[3] http://www.kubios.com/

Wrist Temperature

For analysis, single and dual tasks were compared with the baseline [44]. The *Empatica E4* output temperature values with a sampling rate of four samples per second. Averages were taken for each task for each individual participant.

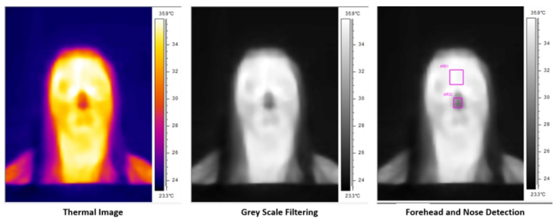


Figure 3: Region of Interest Identification.

Facial Temperature

Using the commercial *FLIR One* thermal camera, we extracted the forehead and nose temperature from the radiometric pictures. We took the mean temperature during each task so that we can study the difference between the increasing number of tasks. we extracted the data of each thermal image we took using *ThermaCam Researcher Professional 2.10* [4]. Firstly, we changed the iron scale into grey scale. Secondly, we put one region on the forehead and the other on the nose using 15×15, 30×30 pixels window on the nose and forehead respectively as shown in Figure 3.

Participants and Procedure

We invited 20 participants (11 females) with an average age of 20.45 years ($SD = 1.14$) using university mailing lists. After arriving in the lab, participants signed a consent form and were briefed with the purpose of the study along with the instructions and descriptions of the experiment. Next, we asked participants to perform the set of single and dual tasks, each for five minutes. Such duration was chosen because HR computation requires data chunks of at least five minutes [41]. The study took approximately 30 minutes (15 minutes tasks and 15 minutes introduction and sensors setup). During the entire experiment, we recorded the participant’s heart rate, heart rate variability, galvanic skin response and skin temperature. The order of the tasks was counter-balanced using balanced Latin-square.

RESULTS

RQ1: Heart Rate as an Indicator of Number of Tasks

In order to answer **RQ1**, we analyzed the effect of the number of tasks on the recorded heart rate as our dependent variable. In this phase we only analyzed the heart rate captured from the *Empatica E4*, as it is the high end version of the heart rate detection methods.

Effect of Number of Tasks on Heart Rate (Empatica E4)

We tested the effect of the TASKS NUMBER on the HEART RATE with a one-way ANOVA. We found a large significant effect of TASKS NUMBER on the HEART RATE mean ($F_{2,38} = 4.66, p < .0.05, ges = 0.20$). Regarding the homogeneity test, variances of the distributions in the population were equal. Also, the dependent variable was normally distributed in each group.

Bonferroni-corrected post-hoc tests found a statistically significant difference between all number of tasks ($p < 0.01$), except between the baseline and single task conditions (Table 1). The mean increase in the heart rate between the increasing number of tasks was of 3.05 (Figure 4).

RQ2: Does it still hold when using the cheap version ?

In order to answer **RQ2**, we analyzed the effect of the number of tasks on the recorded heart rate as our dependent variable. In this phase we analyzed the heart rate captured from the *Polar H6* and compared it to the outcome from the *Empatica E4* sensor. The main aim of this evaluation is to asses the performance of the cheap and affordable heart rate sensor as opposed to the expensive version.

Effect of Number of Tasks on Heart Rate (Polar H6)

We tested the effect of the TASKS NUMBER on the HEART RATE with a one-way ANOVA. Mauchly’s test showed a violation of sphericity against TASKS NUMBER (0.516, p<0.05), so we report Greenhouse-Geisser-corrected ($GGe = 0.67$) values. We found a large significant effect of TASKS NUMBER on the HEART RATE mean ($F_{1.34,25.61} = 8.78, p < 0.005, ges = 0.32$).

Bonferroni-corrected post-hoc tests found a statistically significant difference between the baseline and the dual task conditions ($p < 0.01$), as well as the single and the dual task conditions ($p < 0.05$). However, no significance was found between the baseline and single task conditions (Table 1). The mean increase between the increasing number of tasks measured using the low-cost *Polar H6* chest belt was of 1.675.

Figure 4 shows average HR measurements of both sensors through the different number of tasks. Readings captured by the two sensors were close to each other with maximum difference of 1.88 beats in the baseline phase. This can be due to the differences in algorithms used to calculate HR in each sensor.

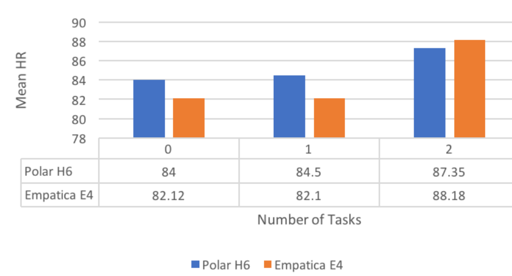


Figure 4: Average HR data from both sensors in different number of tasks

RQ3: Are other signals from commercial sensors capable of distinguishing between the number of tasks?

In order to answer **RQ3**, we analyzed the effect of the number of tasks on the recorded bio-signals. We used five metrics as our dependent variables:

[4]https://thermacam-researcher-pro.software.informer.com/2.1/

1. Increase in Galvanic Skin Response.
2. Variation in Heart Rate Variability components, namely:
   (a) Low Frequency(LF),
   (b) High Frequency(HF),
   (c) Low Frequency to High Frequency (LF/HF) power ratio,
   (d) and RR inter beat intervals.
3. Decrease in wrist temperature.
4. Increase in nose temperature.
5. Increase in forehead temperature.

Effect of Number of Tasks on Galvanic Skin Response

We tested the effect of TASKS NUMBER on the GALVANIC SKIN RESPONSE with a one-way ANOVA. We found no significant effect of TASKS NUMBER on the GALVANIC SKIN RESPONSE (p=0.12).

However, difference was noted upon conducting paired sample t-test between baseline (M=1.2, SD=1.9) and dual task (M=1.5, SD=2.4) conditions; t(19)=-1.81, (p<0.1)(Table 1).

Effect of Number of Tasks on Heart Rate Variability

Tests were administered for LF, HF, RR and LF/HF power ratio using a comparison between baseline, single and dual tasks.

We tested the effect of the TASKS NUMBER on the **HF Power** component with a one-way ANOVA. We found a large significant effect of TASKS NUMBER on the HF POWER component ($F_{2,22} = 9.78, p$<0.001$, ges = 0.47$) (Figure (5a)). Regarding the homogeneity test, variances of the distributions in the population were equal. Also, the dependent variable was normally distributed in each group.
Bonferroni-corrected post-hoc tests found a statistically significant difference between baseline and single task (p<0.05), as well as baseline and dual task conditions (p<0.01). However, no significance was found between single and dual task conditions (Table 1).

Regarding the **LF Power** component, it is suggested that the low frequency component which indicates sympathetic nervous system arousal generally tends to increase in the majority of people upon performing tasks(Figure (5b)). However, ANOVA results showed no significant difference between the three tested states.

Regarding the **(LF/HF) Power Ratio** component, no significant difference was found for the zero-dual nor the single-dual task comparison. However, significant difference was found between the baseline (M=168.27, SD=231.01) compared to the dual task condition(M=25.48, SD=20.61), (p<0.05).

For the **RR** intervals, results showed that 76.92% of the participants experienced a decrease in it during the dual task phase compared to the single phase (Figure 5d). Performing one-way ANOVA test, Mauchly's test showed a violation of sphericity against TASKS NUMBER (0.54, p=0.07), so we report Greenhouse-Geisser-corrected ($GGe = 0.688$) values. We found an effect of TASKS NUMBER on the RR component ($F_{1.3,16.5} = 3.312, p$<0.1$, ges = 0.216$).
Bonferroni-corrected post-hoc tests found no statistically significant difference between all number of tasks, except between the zero and dual task conditions ($p < .05$) (Table 1).

Figure (5c) shows that average HRV power ratio LF/HF power ratio values for the dual task condition decreased noticeably compared to both baseline and single-task phases, with an average difference of 128.42. This shows that the average increase in the HF component was greater than the increase in LF component as shown in Figure 5. Increase in HF component is due to the rapid breath rates induced by stress provoked by the multitasking condition.

Such results further justify the increases noticed in the HF power component, as HF is directly tied to respiration rate. Since RR interval and HR are inversely proportional, RR intervals decreasing means an increase in heart rate. An increase in heart rate means an increase in respiratory rate and hence justifies the increase in the HF values.

Effect of Number of Tasks on Skin Temperature

Wrist Temperature

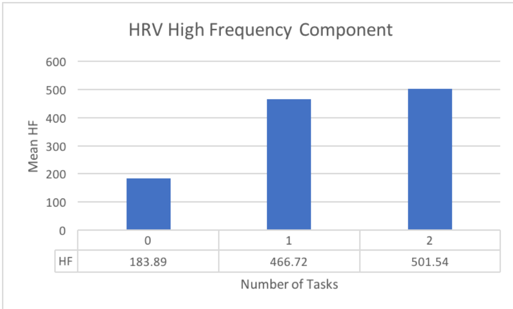
We tested the effect of the TASKS NUMBER on the WRIST TEMPERATURE with a one-way ANOVA. Mauchly's test showed a violation of sphericity against TASK NUMBER (0.34, p<.001), so we report Greenhouse-Geisser-corrected ($GGe = 0.601$) values. We found a large significant effect of TASKS NUMBER on the WRIST TEMPERATURE ($F_{1.2,22.8} = 5.97, p$<0.05$, ges = 0.24$).
Bonferroni-corrected post-hoc tests found a statistically significant difference between both baseline-dual task ($p < 0.01$) and single-dual task comparisons ($p < 0.05$), while baseline-single task conditions comparison showed no significant difference (Table 1). The mean decrease in temperature between increasing number of task was of 0.09 degrees Celsius (Figure 6).

Nose Temperature

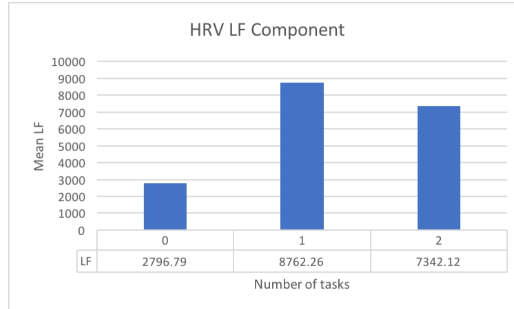
we investigated the effect of the TASK NUMBER on the NOSE TEMPERATURE with a one-way ANOVA. We found no significant effect of TASKS NUMBER on the NOSE TEMPERATURE (p=0.24). Paired sample t-tests found no statistically significant difference between all content types. The mean increase in temperature between increasing number of tasks was of 0.21 degrees Celsius (Figure 6), where 61.9% of the participants experienced an increase in their nose temperature during both single and multitasking activities.

Forehead Temperature

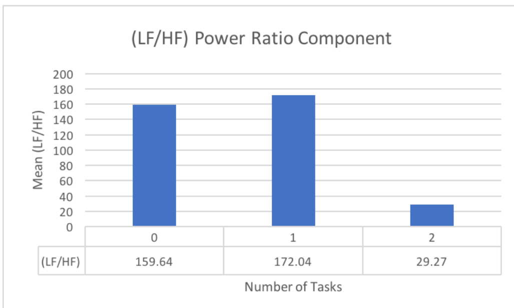
examining the effect of the TASK NUMBER on the FOREHEAD TEMPERATURE with a one-way ANOVA. We found a large significant effect of TASK NUMBER on the FOREHEAD TEMPERATURE ($F_{1.28,34.65} = 12.54, p < 0.001, ges = 0.40$). Regarding the homogeneity test, variances of the distributions in the population were equal. Also, the dependent variable was normally distributed in each group.
Bonferroni-corrected post-hoc tests found a statistically significant difference between all content types ($p < 0.01$), except between single and dual tasks conditions (Table 1). The mean increase in temperature between increasing number of tasks



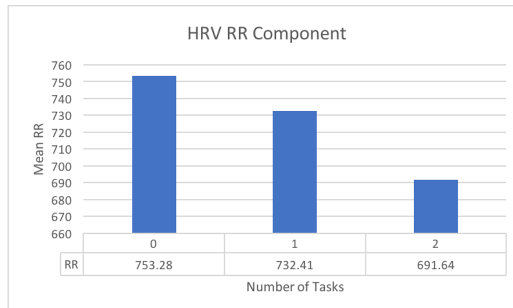
(a) High Frequency



(b) Low Frequency



(c) (LF/HF) Power Ratio



(d) RR Interval

Figure 5: Average of HRV components for different number of tasks

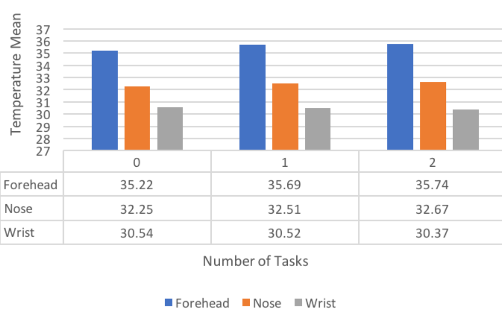


Figure 6: Skin temperature means for different number of tasks

was of 0.36 degrees Celsius (Figure 6), where 76.19% of the participants experienced an increase in their forehead temperature during the single task, and 95.24% experienced an increase in the same area performing the multitasking activity.

Figure 6 shows a comparison between skin temperature of the the three regions of interest (wrist, nose, and forehead) across different number of tasks.

In summary, we found statistically significant effects of task numbers on the following:

1. Heart rate recorded from Empatica E4 relatively expensive commercial sensor ($p = 0.015$)
2. Heart rate recorded from Polar H6 low cost sensor ($p = 0.003$)
3. Hear rate variability HF component ($p = 0.001$)
4. Heart rate variability RR component ($p = 0.07$)
5. Wrist temperature ($p = 0.008$)
6. Forehead temperature from very low cost thermal camera. ($p < 0.001$)

DISCUSSION

Educated by previous work, we conjectured that changing different user states with different number of task assignments would prompt an adjustment in the participants' bio-signals, we wanted to test the possibility of distinguishing between single and multitasking states using relatively low-cost bio-sensors.

Three tasks variations were tested on different metrics: 1) Relaxing as the baseline, 2) Single-task activities, and 3) Dual-task activity. The tested metrics were heart rate, heart rate variability with four tested components (HF, LF, LF/HF power ratio, and RR intervals), galvanic skin response and skin temperature with three regions of interest (wrist, forehead and nose). Our findings were as follows:

Regarding the HEART RATE, readings captured from the *Empatica E4* wristband were firstly analyzed -as it is the high end HR measurement device version- to see whether it can be used as an indicator of number of tasks or not **(RQ1)**. We found a large significant effect of tasks number on the heart rate measurements, where a statistically significant difference was found between all number of tasks, except between the baseline and single task conditions (Table 1).

Afterwards, a first follow up was held through analyzing readings captured by the low-cost heart rate chest belt *(Polar H6)* and comparing them with those of the high end *Empatica E4* sensor with the aim of knowing whether the correlation still holds when using the cheap version of the heart rate sensor **(RQ2)**. The *Polar H6* chest belt results showed that even though it is relatively cheap in comparison to the *Empatica E4* sensor, it can still be used in different task number recognition using HR measurements (Table 1).

		Tasks Number Effect		
Sensor	*Metric*	**0 Vs. 1**	**0 Vs. 2**	**1 Vs. 2**
Polar H6	HR	0.344	0.006***	0.005***
Empatica E4	HR	0.995	0.004***	0.031**
	LF	0.119	0.118	0.773
	HF	0.028**	0.004***	0.419
	(LF/HF)	0.343	0.042**	0.275
	RR	0.241	0.025**	0.209
	GSR	0.162	0.087*	0.199
	Wrist	0.248	0.003***	0.023**
Flir One	Nose	0.144	0.168	0.552
	Forehead	0.002***	0.001***	0.149

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 1: Results Summary.

Based on these results, the rest of the recorded bio-signals were analyzed with the aim of detecting if they are capable of distinguishing between the number of tasks given that they are extracted from affordable, low-end sensors **(RQ3)**.

Accordingly, five metrics were measured: GSR, HRV and wrist temperature using the *Empatica E4* commercially available wristband, in addition to nose and forehead temperature measured by the low-end thermal camera *FLIR One*.

Regarding the *Heart Rate Variability*, upon administering the tests, noticeable increase was found in LF and HF components under single and dual tasks conditions compared to the baseline phase (Figures 5b, 5a). These results suggest that the LF component which indicates sympathetic nervous system arousal generally tends to increase under the increasing number of tasks in the majority of the participants. Additionally, most of the subjects experienced a decrease in the (LF/HF) dual task condition ratio compared to the baseline (Figure 5c). This shows that the average increase in the HF component due to the rapid breath rates was greater than the increase in LF component. Moreover, subjects experienced a decrease in RR intervals during the dual task condition (Figure 5d), which justifies the increases in HF power, as HF is directly tied to respiration rate. Since RR interval and HR are inversely proportional, RR intervals decreasing means an increase in HR, which was confirmed by the HR analysis. To sum up, the results confirm the studies that has considered HRV to be one of the important components in affect recognition along with HR measurements [12, 50, 33, 32].

Regarding the *Galvanic Skin Response*, no significant effect of tasks number was found on the skin conductivity of the participants. However, results suggests that the skin conductance response increases considerably during exposure to the dual task condition compared to the baseline, which confirms the literature reporting that GSR increases with high cognitive load activities [34, 36, 9, 25, 37].

Regarding the skin temperature, three ROIs were investigated being: wrist temperature measured using the *Empatica E4* sensor, along with forehead, nose temperature measured using the low-end *Flir One* thermal camera.

For the *Wrist Temperature*, test results showed a large effect of tasks number on the measurements, where participants experienced a significant decrease in writs temperature with the increasing number of tasks (Figure 6). We found a statistically significant difference between all task types, except between the baseline and single task conditions (Table 1), which comes in line with HR, GSR and HRV analysis results, where no significance was found between the baseline and the single task conditions due to the lack of pressure during the single task.

Concerning the facial temperature, statistically significant difference was found in the *Forehead Temperature* between all task types, except between single and dual tasks conditions (Table 1). The dual tasks made a noticeable change in the forehead temperature as mentioned in previous work [40, 43]. This increase was observed while imposing a sympathetic action and that was reflected in our experiment when 95.24% of participants had their forehead temperature increased under multitasking. It is worth noting that no significant difference in *Nose Temperature* was found neither during the single nor the multitasking conditions. However, most of the participants experienced an increase in the nose temperature with the increasing number of tasks (Figure 6).

Our findings about the increasing facial temperature, and the decreasing wrist temperature with the increasing number of tasks support the literature [37], where hands temperature decreases as a result of blood being shunned away and diverted to more important parts such as the brain, which hence supports the increase in the forehead temperature.

Table 1 summarizes the results of the effect of different tasks number on our measured bio-signals. To sum up, the three research questions were answered using the experimental analysis results as follows:

- **RQ1:** Changes in heart rate measured using a commercially available sensor can distinguish between baseline versus multitasking activities, in addition to single versus multitasking activities.
- **RQ2:** Measuring the heart rate using a cheap, low-end sensor revealed that we can still distinguish the changing nature of the tasks.
- **RQ3:** Some other signals from commercial sensors are capable of accurate distinguishing between the tasks number as well, namely the wrist temperature and the forehead temperature, while other signals, namely the heart rate variability and the skin conductivity showed unified patterns in their change with no significance.

Involvement of Developing Countries

Embracing cross-cultural development, as well as involvement of devolving countries in the research community is of a great importance. While certain research fields are not accessible to developing countries due to the cost of the hardware, we investigate the usage of relatively cheap sensors as an initial step towards their involvement in the affective computing community. This study was held by local researchers with local participants in a newly established research lab of an educational institution in a developing country. This was done with the aim of involving different affect-aware studies and applications in this institution.

Both researchers and participants gave positive feedback about the study, the local researchers reported that they were excited to start working with affordable hardware than can be embedded in their studies for diverse affect-aware applications. Moreover, the participants were interested in exploring the sensors, and using them during the experiment. They reported that they would like to use adaptive systems in their daily life activities given that they are built using user friendly sensors. We envision the involvement of the developing countries would yield the emerging of diverse and tailored set of applications that would be designed, implemented and deployed by and for developing countries.

To sum up, our findings reflect that using affordable methods to meet local requirements of the economically challenged research groups is fruitful, it is for these reasons that we believe it is critical for developing countries to start creating their own research clusters.

LIMITATIONS AND FUTURE WORK

Although our findings suggest the feasibility of using affordable commercial sensors in distinguishing between different number of tasks using different bio-signals measurements, our approach have its own limitations. Single-task activities might not have been challenging enough for the participants to get aroused; this explains the analysis showing that there was no significant difference noted between baseline and single-task conditions in most of the bio-signals. Accordingly, it is recommended to hold further studies that explore more engaging activities. Additionally, extra validation is needed for the GSR, and HRV signals either from other sensors or other experimental conditions to reassure their capability of distinguishing between the increasing number of tasks. Moreover, we used an artificial task to elicit multitasking. It would be interesting to investigate the effect of naturalistic multitasking task on the measured bio-signals.

In future work, our results can be used to generalize the effect of different states on bio-signals using different commercial sensors. Wider scenarios can be considered including different number of tasks and mental load levels with various natures and difficulties in specific contexts like education. Additionally, difference between dominant and non-dominant side measurements of the wristband can be investigated, along with using cheaper sensors for different signals and comparing them with their advanced counterparts. Regarding thermal imaging, more ROIs other than forehead and nose can be considered. Furthermore, this work can be extended to include indicators that measure both arousal and valence like eye and brain activities to insure the accuracy of the results, and embed the outcomes into machine learning systems that detect and predict levels of difficulty faced by the users.

Moreover, our work highlight the potential of using biosensors to distinguish between number of tasks being performed. We hope our work could inform future research to build a model to investigate with which accuracy the number of tasks can be differentiated.

CONCLUSION

Nowadays multitasking is heavily included in our daily life routine in different contexts. Ubiquitous computing sensors and affect-aware systems have the capability to support users during performing their tasks through sensing their internal states and adapting their behavior accordingly. Hence, a clear understanding of the effect of multitasking on the user's affect is needed in order to build adaptive systems that interpret and consider the number of tasks in hands.

This study questioned the ability of using commercial biosensors to estimate the state of the user during multitasking itself rather than the induced changes. We investigated the impact of number of tasks on different signals namely HR, HRV, GSR and skin temperature. A set of unobtrusive commercial sensors was chosen so that they can be used during the day, and not just be limited to lab usage. The tasks variations were as follows: Relaxing as the baseline, Single-task and Dual-task activities.

A preliminary study was conducted on HR measurements using a high-end commercial sensor. This was followed by a first follow-up exploration in which we compared the results of this sensor with those of a low-cost chest belt with the aim of validating the usage of affordable sensors to be useful for all groups in developing countries. Findings showed that changes measured using the commercial HR sensor can distinguish between baseline versus multitasking, and single versus multitasking activities. Using the low-end HR sensor revealed that the correlation still holds with the changing nature of the tasks.

A second follow-up exploration with the aim of testing other signals from commercial sensors showed that they are capable of accurately distinguishing between user's states performing different number of tasks, namely wrist and forehead temperature; while other signals like HRV and GSR showed unified patterns in their change with no significance.

Our results can be used with the adaptive systems currently being established in developing countries such as mood based applications that help those who might struggle in expressing themselves during multitasking based workload.

ACKNOWLEDGMENT

This work was supported by the German Federal Ministry of Education and Research, FeuerWeRR Grant No. 13N13481. This work was partly conducted within the Amplify project which received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 683008) and received funding from the German Research Foundation within the SimTech Cluster of Excellence (EXC 310/2).

REFERENCES

1. Yomna Abdelrahman, Mohamed Khamis, Stefan Schneegass, and Florian Alt. 2017a. Stay cool! understanding thermal attacks on mobile-based user authentication. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 3751–3763.
2. Yomna Abdelrahman, Pascal Knierim, Pawel W. Wozniak, Niels Henze, and Albrecht Schmidt. 2017b. See Through the Fire: Evaluating the Augmentation of Visual Perception of Firefighters Using Depth and Thermal Cameras. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers (UbiComp '17)*. ACM, New York, NY, USA, 693–696. DOI:http://dx.doi.org/10.1145/3123024.3129269
3. Yomna Abdelrahman, Alireza Sahami Shirazi, Niels Henze, and Albrecht Schmidt. 2015. Investigation of Material Properties for Thermal Imaging-Based Interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 15–18. DOI: http://dx.doi.org/10.1145/2702123.2702290
4. Yomna Abdelrahman, Eduardo Velloso, Tilman Dingler, Albrecht Schmidt, and Frank Vetere. 2017. Cognitive Heat: Exploring the Usage of Thermal Imaging to Unobtrusively Estimate Cognitive Load. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 33 (2017), 20.
5. U Rajendra Acharya, K Paul Joseph, Natarajan Kannathal, Choo Min Lim, and Jasjit S Suri. 2006. Heart rate variability: a review. *Medical and biological engineering and computing* 44, 12 (2006), 1031–1051.
6. Jenni Anttonen and Veikko Surakka. 2005. Emotions and heart rate while sitting on a chair. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 491–499.
7. Jorn Bakker, Mykola Pechenizkiy, and Natalia Sidorova. 2011. What's your current stress level? Detection of stress patterns from GSR sensor data. In *Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on*. IEEE, 573–580.
8. Wolfram Boucsein. 2012. *Electrodermal activity*. Springer Science & Business Media.
9. Frédéric Bousefsaf, Choubeila Maaoui, and Alain Pruski. 2014. Remote detection of mental workload changes using cardiac parameters assessed with a low-cost webcam. *Computers in biology and medicine* 53 (2014), 154–163.
10. Jason J Braithwaite, Derrick G Watson, Robert Jones, and Mickey Rowe. 2013. A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology* 49 (2013), 1017–1034.
11. Paul W Burgess, Emma Veitch, Angela de Lacy Costello, and Tim Shallice. 2000. The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia* 38, 6 (2000), 848–863.
12. Adrian Burns, Barry R Greene, Michael J McGrath, Terrance J O'Shea, Benjamin Kuris, Steven M Ayer, Florin Stroiescu, and Victor Cionca. 2010. SHIMMERâĎć–A wireless sensor platform for noninvasive biomedical research. *IEEE Sensors Journal* 10, 9 (2010), 1527–1534.
13. TaskForceoftheEuropean Societyof Cardiology. 1996. the North American SocietyofPacingandElectrophysiology. *Heart rate variability: Standards of measurement, physiologicalinterpretation, andclinicaluse. EurHeartJ* 17, 3 (1996), 354–81.
14. Jongyoon Choi and Ricardo Gutierrez-Osuna. 2009. Using heart rate monitors to detect mental stress. In *Wearable and Implantable Body Sensor Networks, 2009. BSN 2009. Sixth International Workshop on*. IEEE, 219–223.
15. Gari D Clifford. 2002. *Signal processing methods for heart rate variability*. Ph.D. Dissertation. University of Oxford Oxford.
16. Roberto Colom, Agustín Martínez-Molina, Pei Chun Shih, and José Santacreu. 2010. Intelligence, working memory, and multitasking performance. *Intelligence* 38, 6 (2010), 543–551.
17. Roddy Cowie, Ellen Douglas-Cowie, Nicolas Tsapatsoulis, George Votsis, Stefanos Kollias, Winfried Fellenz, and John G Taylor. 2001. Emotion recognition in human-computer interaction. *IEEE Signal processing magazine* 18, 1 (2001), 32–80.
18. Iwona Cygankiewicz and Wojciech Zareba. 2013. Heart rate variability. *Handb Clin Neurol* 117, 3 (2013), 379–393.
19. Mai ElKomy, Yomna Abdelrahman, Markus Funk, Tilman Dingler, Albrecht Schmidt, and Slim Abdennadher. 2017. ABBAS: An Adaptive Bio-sensors Based Assistive System. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2543–2550.
20. Markus Funk, Tilman Dingler, Jennifer Cooper, and Albrecht Schmidt. 2015. Stop helping me-I'm bored!: why assembly assistance needs to be adaptive. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*. ACM, 1269–1273.
21. Polar Global. 2017. H6 heart rate sensor product support. (2017). http://support.polar.com/en/support/H6_heart_rate_sensor.

22. Charles J Golden and Shawna M Freshwater. 1978. Stroop color and word test. *age* 15 (1978), 90.
23. R Hamed, Adham Atyabi, Antti Rantanen, Seppo J Laukka, Samia Nefti-Meziani, Janne Heikkilä, and others. 2015. Predicting the valence of a scene from observersâĂŹ eye movements. *PloS one* 10, 9 (2015), e0138198.
24. Khaled Kassem, Jailan Salah, Yasmeen Abdrabou, Mahesty Morsy, Reem ElGendy, Yomna Abdelrahman, and Slim Abdennadher. 2017. DiVA: exploring the usage of pupil diameter to elicit valence and arousal. In *Proceedings of the 16th International Conference on MUM. ACM*.
25. Ahmad Khawaji, Jianlong Zhou, Fang Chen, and Nadine Marcus. 2015. Using galvanic skin response (GSR) to measure trust and cognitive load in the text-chat environment. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 1989–1994.
26. Robert E Kleiger, Phyllis K Stein, and J Thomas Bigger. 2005. Heart rate variability: measurement and clinical utility. *Annals of Noninvasive Electrocardiology* 10, 1 (2005), 88–101.
27. Thomas Kosch, Yomna Abdelrahman, Markus Funk, and Albrecht Schmidt. 2017. One size does not fit all: challenges of providing interactive worker assistance in industrial settings. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*. ACM, 1006–1011.
28. Jure Kranjec, S Beguš, G Geršak, and J Drnovšek. 2014. Non-contact heart rate and heart rate variability measurements: A review. *Biomedical Signal Processing and Control* 13 (2014), 102–112.
29. Peter J Lang. 1995. The emotion probe: Studies of motivation and attention. *American psychologist* 50, 5 (1995), 372.
30. Lars Lischke, Sven Mayer, Katrin Wolf, Niels Henze, Harald Reiterer, and Albrecht Schmidt. 2016. Screen arrangements and interaction areas for large display work places. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays*. ACM, 228–234.
31. Federico Lombardi. 2004. Physiological understanding of HRV components. *Dynamic Electrocardiography* 40, 7 (2004).
32. Daniel McDuff, Sarah Gontarek, and Rosalind Picard. 2014. Remote measurement of cognitive stress via heart rate variability. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE, 2957–2960.
33. Paolo Melillo, Marcello Bracale, and Leandro Pecchia. 2011. Nonlinear Heart Rate Variability features for real-life stress detection. Case study: students under stress due to university examination. *Biomedical engineering online* 10, 1 (2011), 96.
34. Arturo Nakasone, Helmut Prendinger, and Mitsuru Ishizuka. 2005. Emotion recognition from electromyography and skin conductance. In *Proc. of the 5th International Workshop on Biosignal Interpretation*. 219–222.
35. DA Norman and SW Draper. 1986. User Centered System Design: New Perspectives on Human-Computer Interaction Lawrence Erlbaum Associates. (1986).
36. Nargess Nourbakhsh, Yang Wang, Fang Chen, and Rafael A Calvo. 2012. Using galvanic skin response for cognitive load measurement in arithmetic and reading tasks. In *Proceedings of the 24th Australian Computer-Human Interaction Conference*. ACM, 420–423.
37. Simon Ollander. 2015. Wearable sensor data fusion for human stress estimation. (2015).
38. Jani Penttilä, Antti Helminen, Tuomas Jartti, Tom Kuusela, Heikki V Huikuri, Mikko P Tulppo, René Coffeng, and Harry Scheinin. 2001. Time domain, geometrical and frequency domain analysis of cardiac vagal outflow: effects of various respiratory patterns. *Clinical Physiology* 21, 3 (2001), 365–376.
39. Bruce Pomeranz, RJ Macaulay, Margaret A Caudill, Ilan Kutz, Dan Adam, DAVID Gordon, Kenneth M Kilborn, A Clifford Barger, Daniel C Shannon, Richard J Cohen, and others. 1985. Assessment of autonomic function in humans by heart rate spectral analysis. *American Journal of Physiology-Heart and Circulatory Physiology* 248, 1 (1985), H151–H153.
40. Colin Puri, Leslie Olson, Ioannis Pavlidis, James Levine, and Justin Starren. 2005. StressCam: non-contact measurement of users' emotional states through thermal imaging. In *CHI'05 extended abstracts on Human factors in computing systems*. ACM, 1725–1728.
41. Mario Salai, István Vassányi, and István Kósa. 2016. Stress detection using low cost heart rate sensors. *Journal of healthcare engineering* 2016 (2016).
42. Nainal Sanjay Shah. 2008. *Haemodynamic parameters in hypertensive patients following acute stroke and early outcome*. Ph.D. Dissertation. University of Leicester (United Kingdom).
43. Dvijesh Shastri, Avinash Wesley, and Ioannis Pavlidis. 2008a. Contact-free Stress Monitoring for User's Divided Attention. In *Human Computer Interaction*. InTech.
44. Dvijesh Shastri, Avinash Wesley, and Ioannis Pavlidis. 2008b. Contact-free Stress Monitoring for User's Divided Attention. In *Human Computer Interaction*. InTech.
45. Yu Shi, Natalie Ruiz, Ronnie Taib, Eric Choi, and Fang Chen. 2007. Galvanic skin response (GSR) as an index of cognitive load. In *CHI'07 extended abstracts on Human factors in computing systems*. ACM, 2651–2656.

46. Shiv Naresh Shivhare and Saritha Khethawat. 2012. Emotion detection from text. *arXiv preprint arXiv:1205.4944* (2012).

47. Erin Treacy Solovey, Francine Lalooses, Krysta Chauncey, Douglas Weaver, Margarita Parasi, Matthias Scheutz, Angelo Sassaroli, Sergio Fantini, Paul Schermerhorn, Audrey Girouard, and others. 2011. Sensing cognitive multitasking for a brain-based adaptive user interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 383–392.

48. Empatica Store. 2017. E4 wristband Rev.2. (2017). http://support.polar.com/en/support/H6_heart_rate_sensor.

49. Feng-Tso Sun, Cynthia Kuo, Heng-Tze Cheng, Senaka Buthpitiya, Patricia Collins, and Martin Griss. 2010. Activity-aware mental stress detection using physiological sensors. In *International Conference on Mobile Computing, Applications, and Services*. Springer, 282–301.

50. Joachim Taelman, Steven Vandeput, Arthur Spaepen, and Sabine Van Huffel. 2009. Influence of mental stress on heart rate and heart rate variability. In *4th European conference of the international federation for medical and biological engineering*. Springer, 1366–1369.

51. Egon L Van den Broek. 2013. Ubiquitous emotion-aware computing. *Personal and Ubiquitous Computing* 17, 1 (2013), 53–67.

52. Murlikrishna Viswanathan, Zhen-Xing Zhang, Xue-Wei Tian, and Joon S Lim. 2011. A fatigue detection algorithm by heart rate variability based on a neuro-fuzzy network. In *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication*. ACM, 83.

53. Alexandra Voit, Benjamin Poppinga, Dominik Weber, Matthias Böhmer, Niels Henze, Sven Gehring, Tadashi Okoshi, and Veljko Pejovic. 2016. UbiTtention: Smart & Ambient Notification and Attention Management. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*. ACM, New York, NY, USA, 1520–1523. DOI: http://dx.doi.org/10.1145/2968219.2968542

54. Tanja GM Vrijkotte, Lorenz JP Van Doornen, and Eco JC De Geus. 2000. Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability. *Hypertension* 35, 4 (2000), 880–886.

55. Julia Wache. 2014. The secret language of our body: Affect and personality recognition using physiological signals. In *Proceedings of the 16th International Conference on Multimodal Interaction*. ACM, 389–393.

56. Dominik Weber, Alireza Sahami Shirazi, Sven Gehring, Niels Henze, Benjamin Poppinga, Martin Pielot, and Tadashi Okoshi. 2016. Smarttention, Please!: 2Nd Workshop on Intelligent Attention Management on Mobile Devices. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '16)*. ACM, New York, NY, USA, 914–917. DOI: http://dx.doi.org/10.1145/2957265.2965025