

Research Article

Making Wearable Technology Available for Mental Healthcare through an Online Platform with Stress Detection Algorithms: The Carewear Project

Glen Debard ¹, Nele De Witte,² Romy Sels,¹ Marc Mertens,^{1,3} Tom Van Daele,² and Bert Bonroy¹

¹Mobilab & Care, Thomas More University of Applied Sciences Kempen, Geel, Belgium

²Expertise Unit Psychology, Technology & Society, Thomas More University of Applied Sciences, Antwerp, Belgium

³KU Leuven, Department of Computer Science, DTAI, Leuven, Belgium

Correspondence should be addressed to Glen Debard; glen.debard@thomasmore.be

Received 27 March 2020; Revised 20 October 2020; Accepted 28 October 2020; Published 25 November 2020

Academic Editor: Christos Riziotis

Copyright © 2020 Glen Debard et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Over the past years, mobile health (mHealth) applications and specifically wearables have become able and available to collect data of increasing quality of relevance for mental health. Despite the large potential of wearable technology, mental healthcare professionals are currently lacking tools and knowledge to properly implement and make use of this technology in practice. The Carewear project is aimed at developing and evaluating an online platform, allowing healthcare professionals to use data from wearables in their clinical practice. Carewear implements data collection through self-tracking, which is aimed at helping people in their behavioral change process, as a component of a broader intervention or therapy guided by a mental healthcare professional. The Empatica E4 wearables are used to collect accelerometer data, electrodermal activity (EDA), and blood volume pulse (BVP) in real life. This data is uploaded to the Carewear platform where algorithms calculate moments of acute stress, average resting heart rate (HR), HR variability (HRV), step count, active periods, and total active minutes. The detected moments of acute stress can be annotated to indicate whether they are associated with a negative feeling of stress. Also, the mood of the day can be elaborated on. The online platform presents this information in a structured way to both the client and their mental healthcare professional. The goal of the current study was a first assessment of the accuracy of the algorithms in real life through comparisons with comprehensive annotated data in a small sample of five healthy participants without known stress-related complaints. Additionally, we assessed the usability of the application through user reports concerning their experiences with the wearable and online platform. While the current study shows that a substantial amount of false positives are detected in a healthy sample and that usability could be improved, the concept of a user-friendly platform to combine physiological data with self-report to inform on stress and mental health is viewed positively in our pilots.

1. Introduction

Over the past years, mobile health (mHealth) applications and specifically wearables have become able and available to collect data of increasing quality that is of relevance for mental health. Wearables allow the continuous and ecologically valid collection of physiological data and can, furthermore, obtain relevant information at different stages of mental health disorders: from initial risk factors, over treatment progress, all the way to the process of recovery [1].

However, several challenges remain to make this sensor data useful and usable for mental healthcare in real life. Substantial technical capacity is needed for data handling and analysis. Additionally, data needs to be made available for use by mental healthcare professionals and clients in a collaborative space, where it becomes actionable and interpretable. Despite the large potential of wearable technology, mental healthcare professionals are currently lacking the tools and the knowledge to properly implement and make use of this technology in real life [2].

Wearables are a specific type of mHealth application consisting of sensors and devices that can be worn on the body and can collect longitudinal and continuous data on a reliable and noninvasive manner outside of lab settings. However, Can et al. [3], Larradet et al. [4], and Sun et al. [5] state that measuring physiological signals during everyday activity is more difficult than in lab conditions. A first challenge is that the physiological responses of mental stress can be masked by physical activity. Secondly, the accuracy of the measurements is affected by signal artefacts caused by motion, electrode placement, or respiratory movement. Thirdly, for training a stress model, it is difficult to determine the user's stress level in natural circumstances to label the training data [6]. Additionally, the stress level that is then determined through self-report is the perceived stress level [7], which might be different from their physiological stress level. Kyriakou et al. [8] observed that sometimes a physiological stress state was not recognized by the participants of their research. The self-reported stress moments can also be shaped by many social factors, leading to discordance between physiological and self-reported stress or emotional measures.

Gaining information about relevant parameters in daily life, such as stress and physical activity, has always been essential for tailoring and evaluating interventions in psychotherapy and counseling. Such data is collected through self-report information in sessions or between sessions, through pen-and-paper diary methods, or more recently also through computer or smartphone apps. Asking clients to report on events and emotions during the past week(s) has the potential limitation that it can be subjected to memory or reporting bias. For example, individuals, especially those with a history of depression, appear to overestimate the daily occurrence of negative emotions [9]. Additionally, people have a tendency to forget emotional peaks within 24 hours [7]. Diary methods have been used successfully in treatment and research, but often show decreased use over time [10, 11]. Wearables could help collect data between sessions that is less subject to bias and such automatic registration could reduce the load and hereby potentially increase adherence to data collection between sessions. The study of Patel et al. [12] shows that most individuals show continued use of wearable data for over 6 months.

The review of Kersten-van Dijk and colleagues [13] provides evidence that personal informatics, which refers to using technological devices to monitor and review personally relevant data, can provide end users with new insights and raise awareness about stress for example. However, it is key that the data is actionable and that sufficient support is provided. Platforms that make wearable data available for research have already been developed [14], but do not support clinical application.

Different types of wearables are used in the literature on stress detection. For example, Sun et al. [5] and Han et al. [15] used Shimmer sensors that can be placed on different parts of the body; Hovsepian et al. [16] and Rahman et al. [17] used the AutoSense sensor suite that consists of a flexible band worn around the chest. Tazawa et al. [18] used a Silmee W20 wristband, de Arriba-Pérez et al. [19] used the (now discontinued) Microsoft Band 2 wrist wearable, and Mishra

et al. [20] used the commercially available Polar H7 chest heart sensor. Other studies used a combination of different sensors that are difficult to use in real life [21–23]. Nevertheless the wristband is the most common example of a wearable suitable for real-life measurement and has already been used in different settings over time. Developments include the use of smartwatch sensory data to measure indicators of mental health in schizophrenia [24], the application of accelerometer data as a biomarker for depression [25, 26], and the design of a smartwatch application for the management of ADHD [27]. Wearable technology can provide additional data for the prevention and treatment of disorders, map the effects of interventions, and provide momentary feedback. However, careful selection of a manageable set of physiological and behavioral parameters of interest for mental healthcare is important.

Wearable monitors can collect data on cardiac cycles, electrodermal activity (EDA), skin temperature (ST), and acceleration. Momentary increases in heart rate (HR) and EDA, which was checked for physical activity, could be indicative of stress. Stress consists of a complex interplay between psychological, behavioral, and physiological responses evoked by a psychological or physical threat to homeostasis [28]. Changes in HR, blood pressure, EDA, and breathing rate are commonly observed in stressful situations [29–31] and could therefore aid in stress detection. Nevertheless, self-report information remains important as well, since deriving valence from physiological activation is very difficult. Moreover, stress detections only become relevant for psychological prevention and therapy when context is provided.

Two other relevant parameters that can be calculated are physical and sport activities and heart rate variability (HRV). Stimulating physical activity has been shown to reduce depressive symptoms and stress levels [32–34] and promote recovery from burnout, depression, and anxiety [35]. Finally, HRV refers to variations in beat-to-beat intervals controlled by the parasympathetic nervous system and prefrontal cortex. Previous research has stated that HRV is an index of flexibility to cope with complex challenges, and low HRV can be a sign of chronic stress and allostatic load [36, 37]. Elevated stress, burnout, and depression are associated with reduced HRV [37–39]. Wearables can monitor resting HRV, which could potentially inform on resilience or risk for a mental illness. However, since wearable devices with an adequate sampling rate for HRV calculation are only being developed recently, there is a lack of longitudinal ambulatory HRV monitoring studies. The validity of such a longitudinal HRV assessment is, therefore, still to be determined.

To increase the odds of continuous stress detection using wearables in real-life settings, as opposed to lab and research settings, it is important to opt for commercially available devices that are easy to use and wear. Gradl et al. [40] gave an overview of existing wearables together with their measurable parameters. They also rated each wearable with their estimated potential to measure stress. They rated wearables that are able to measure EDA, such as the Empatica E4 and Sentio Feel, highest.

Most studies in the literature that try to determine stress levels use a combination of several parameters extracted from one or more physiological signals. Features extracted from EDA are used most often [5, 6, 8, 15, 22, 23, 41–47], closely followed by features extracted from the cardiac cycle [5, 6, 15, 16, 21–23, 41, 42, 44, 46, 47]. These last studies mainly use an ECG monitor in different forms to extract the signal [5, 6, 15, 16, 21–23, 41, 42]. However, such a monitor is more difficult to use outside of the lab as also stated by Rahman et al. [17], so for real life, devices using photoplethysmography (PPG) to measure BVP are preferred. Additionally, respiration rate [15, 16, 22, 23, 41] and skin temperature (ST) [6, 8, 22, 43, 46] are regularly used.

Larradet et al. [4] wrote an extensive review in which they presented the main differences between classification and detection of stress and emotions according to data collected in real life or in the laboratory. They state that EDA, ECG, and EMG can greatly differ between real-life and laboratory settings. So there is a real need for research to be done in emotions recognition in real life. They showed that, while there has been some research in this area, there are still very few papers focusing on this matter today.

As stated above, self-reports are mostly used in real-life experiments to define the ground truth. Because of this, Can et al. [3] state that achieving precise annotations and identification of the perceived stress in real life is a difficult task. They also state that the stress level experienced in the laboratory is different from daily life stress. Because of this, they conclude that using a model that is trained using laboratory data to classify real-life events outperforms a model that is solely trained using real-life data.

The resulting information from the different stress detection studies also varies in purpose. Sometimes the stressful periods over the course of the day are detected [8, 16, 22, 41, 46], and in other cases, the stress level of an event or period, lasting, e.g., twenty minutes, is registered [15, 41, 43, 44, 46]. Another approach is to define the overall stress level of a day [21].

Previous diary-based research has suggested that having more data points per day, as opposed to general daily stress reports, will be better able to capture the relationship between stress and behavior [11]. While these authors suggested that increased reporting could lead to increased burden and decreased willingness in participants, wearable monitoring might nevertheless facilitate data sampling hence decreasing burden. Thus, providing insight into specific moments of stress during the day, by integrating real-life stress detections with additional contextual self-report data, could provide the healthcare professional and their client with actionable data allowing to uncover patterns and tailor interventions. Because of this, we want to detect short moments of acute stress caused by the most stressful events throughout the day to give the user the opportunity to annotate these moments and discuss them with their mental healthcare professional.

The Carewear project (HBC.2016.0099), therefore, is aimed at developing and evaluating an online platform, allowing healthcare professionals to use data from wearables in their clinical practice. Carewear implements data collec-

tion through self-tracking, with the aim to help people in their behavioral change process, but only as a component of a broader intervention guided by a mental healthcare professional. The online platform provides aggregated variables and allows to integrate wearable data with personal experience. A psychologist of the team (NDW) also developed two accompanying manuals [48, 49] for the online platform to support the users of the platform. The first manual is focused on practical information for both professional and end user, consisting of how to wear the wearable and handle the data. The second manual is for professionals only and provides practice-oriented information on how to use the platform in an evidence-based way in the context of stress-related complaints and depression.

As also stated by Can et al. [3], Larradet et al. [4], and Sun et al. [5], detecting stress in real life is much more difficult than in lab conditions. Because of this, the current study's goal was a first assessment of the performance of the implemented algorithms through comparisons with comprehensive annotated data in a small healthy sample captured in real life. Additionally, we assessed the usability of the application through reports of the users on their experiences using the wearable and online platform.

2. Materials and Methods

The Empatica E4 [50] wearables were used for data collection (see Figure 1). This wristband is a class IIa medical device and can collect accelerometer (ACC) data, electrodermal activity (EDA), skin temperature (ST), and blood volume pulse (BVP). Clients collect data with this wristband and afterwards upload it to the Carewear platform. Algorithms are implemented to remove artefacts and transform the raw data into interpretable indicators, consisting of acute stress moments, step count, minutes of increased physical activity (sports), mean HR, and HRV (see Section 2.4). Users can consequently consult and complete their data using the Carewear platform typically once each day. A physiological stress detection that is classified as an acute stress moment will only be logged as an actual stressful event after the client has verified this moment on the online platform (refer to Section 2.3). Providing additional, contextual information to the data allows the client and healthcare professional to discover patterns and tailor interventions to the actual needs.

2.1. Wearable. As mentioned above, the Empatica E4 integrates different sensors in one wrist-worn wearable. Menghini et al. [51] validated the accuracy of this device. They concluded that it provided an accurate mean HR in both static and dynamic conditions. HRV was accurate in static conditions. The accuracy was less reliable in hand movement conditions. The accelerometer is a 3-axis accelerometer sensor that works in the range of -2 g till 2 g. The sample frequency is 32 Hz. The BVP data is collected using photoplethysmograph (PPG). This sensor uses green and red lights that are reflected as a function of the blood oxygenation. The more the blood is oxygenated, the more the light is absorbed. Thus, during a heartbeat, less light is reflected. The sample frequency of the BVP signal is 64 Hz. The Empatica



FIGURE 1: The Empatica E4.

E4 captures the EDA by measuring the electrical conductance across the skin. It achieves this by passing a minuscule amount of current between two electrodes in contact with the skin on the inside of the wrist. The data from the EDA sensor is sampled at 4 Hz. The wearable can also detect the skin temperature, but this parameter is not included in the current study. Finally, the user can press the operating button on the Empatica E4 whenever an event occurs which the user wants to manually register to discuss with their mental healthcare professional. This manually tagged event is subsequently shown in the platform to be annotated. However, for this study, we asked the participants not to use this feature.

2.2. Sample. Participants were recruited from a healthy student sample. The sample consists of four female and two male participants with a mean age of 20.5 years old ($SD = 0.8$). Participants wore the Empatica E4 on their non-dominant hand, which was the left hand for all participants, to reduce the risk of movement artefacts for about one week during all daily activities. They also kept a detailed journal of activities and stress-related events. One included participant had a diagnosis of attention deficit hyperactivity disorder (ADHD) and reported the use of methylphenidate hydrochloride in their detailed journal. Participant 3 used the wearable for several days, but did not annotate the detected moments of acute stress, so this person had to be excluded. So the final sample consists of five participants. The study was approved by the ethical committee of the Department of Applied Psychology of Thomas More University of Applied Sciences, and all participants provided informed consent.

We created a table with relevant labels for the data logging during the measurement period. The participants were asked to continuously report on the activity they were performing (e.g., eating, following class, and jogging) and on whether they experienced any increased arousal. Additionally, they were asked to keep note of every stressful occur-

rence, in their opinion, they experienced so as to compare it with the data in the online platform. The listed stressful events contained, but are not limited to, presenting before audience, examinations, running late for public transportation or another appointment, driving a car, or being startled because of remembering something they forgot or something falling over or on the ground. At the end of the period, they provided a written report detailing their personal experiences with the wearable and platform.

The participants were given the necessary materials for participation, consisting of a wearable, access to the online platform, and documents to provide a detailed report on their activities. They collected and annotated data during 1 week. Afterwards, they also provided a written report of their experiences with the wearable and online platform.

2.3. Online Platform. An online platform was developed for the Carewear project. Both the client and their mental healthcare professional each have their own interface.

- (1) *Development process of the platform:* the online Carewear platform was designed in close collaboration with organizations that were either offering employee assistance programs, were specialized in technological development, or provided clinically oriented services. These organizations provided input and feedback for the development of the online platform during several user group meetings and discussions. The design of the platform was subsequently further scrutinized from the perspective of mental healthcare professionals in a focus group of five lecturers in applied psychology, of which three male and two female. Each had expertise in the practice of mental healthcare and was aware of research in this domain and new trends in this regard. The received feedback included, but was not limited to, warranting care in using the colors green and red, as they can trigger a

negative emotion, use a slider instead of only the dichotomous options positive or negative to annotate how a user felt during a given stress event, and to assure user privacy. The perspective of potential end users was also covered in a focus group with six female final-year students in applied psychology. The end users found it for example important to have a means to provide feedback of how they felt over the complete day and to have a clear view of their goals. The input of these different groups were integrated to create the first working prototype of the online platform.

This first prototype of the Carewear platform was then implemented in a pilot with five professionals: two mental healthcare practitioners, two professionals offering employee assistance programs, and one representative of the Regional Federation of Psychological Consultants. Participants were generally favorable towards the concept of making noninvasively collected continuous physiological data available as an additional source of information that can contribute to better, personalized care. The positive aspects, which had them believe this could actually be of added value to treatment as usual, concerned the fact that Carewear coupled a technical solution with a manual for clinical application. Reported negative aspects consisted of difficulties with uploading data (which were related to the absence of an Empatica rest interface), measurement errors that occurred which did not seem to reflect their real-life experiences, no feedback when pressing twice for HRV measurement, and usability aspects that could be improved. One professional was less interested in applying this platform for stress-related complaints, but reported that “such an application would be very useful for people who have difficulty connecting with what is happening in their body” and was interested in using the platform in chronic fatigue or pain patients. Another clinician was also interested in using Carewear in the context of anxiety disorders, which is nevertheless out of the scope of this work. Overall, these results were used to update the Carewear platform in terms of usability.

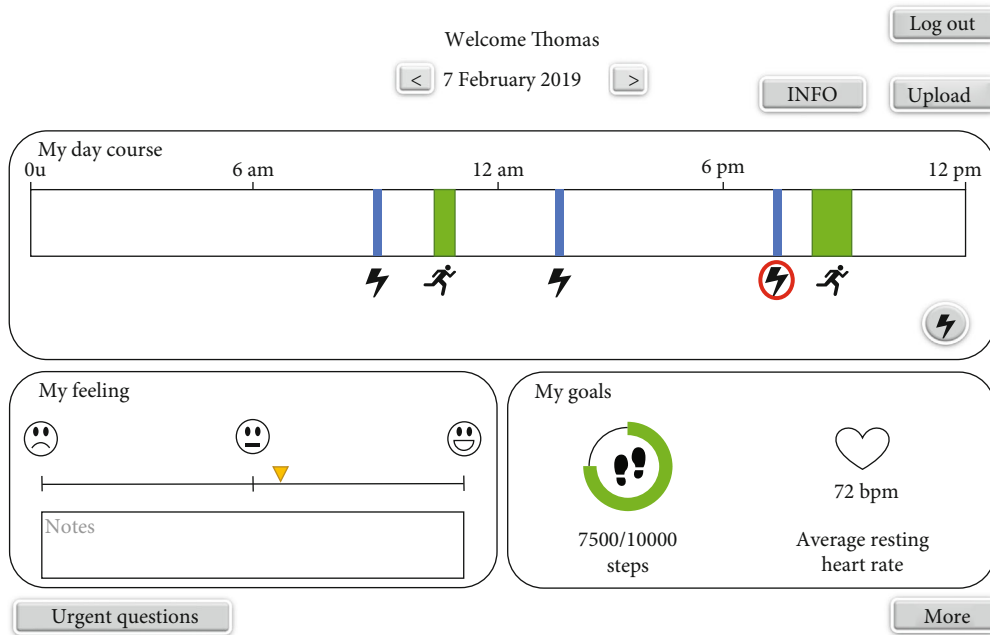
- (2) *Features of the implemented platform:* the Carewear platform from the client’s point of view consists of a home page with an overview of the current day, from which they can navigate to a detailed view of the day, and a weekly and monthly overview as shown in Figure 2. The home page (see Figure 2(a)) shows an overview of their day containing the active periods, the button-pressed events, and the detected moments of acute stress. For each of these stress events, the user can fill in additional information on their subjective experience related to this event through a pop-up screen (see Figure 3). In case the user remembers the moment of the detected event, they can indicate whether the subjective experience was associated with a negative feeling of stress or not. Further

information about how they felt at that time, what was the cause, and how long the feeling lasted can be entered. The overall mood of that day with a possible clarification can also be filled in. Additionally, the total step count and the average resting heart rate of that day are shown.

The detailed view of the day (see Figure 2(b)) shows the step count per hour, which is a downsampled plot of the EDA and HR combined with an indication of the stress detections, total step count, average resting HR, total active minutes, and mood of the day. The weekly and monthly overviews (see Figure 2(c)) show a graph of the total step count, active minutes, overall mood, number of stress detections, and HRV with a linear regression line of its evolution plotted per day.

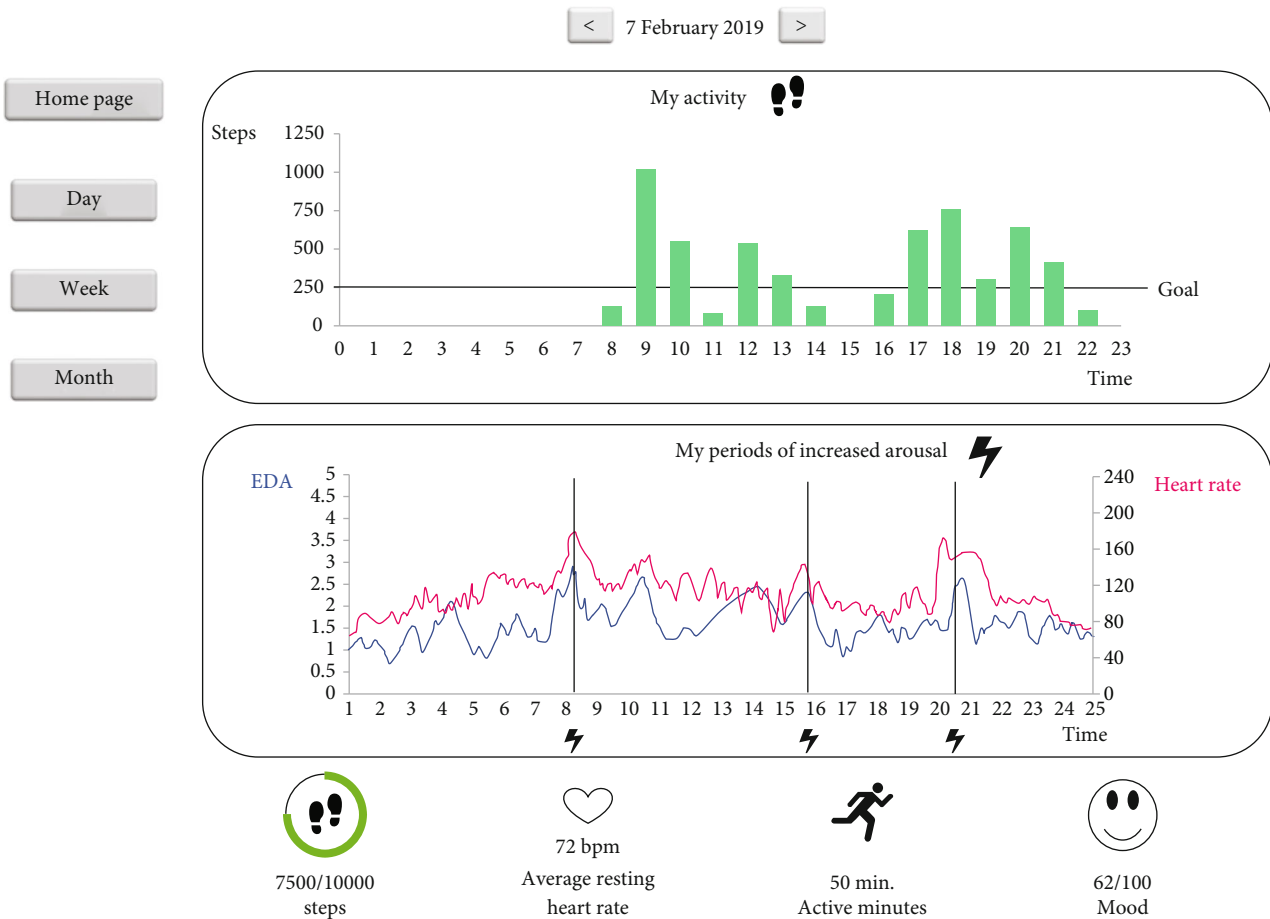
The mental healthcare professional can also view the client’s pages and, additionally, has a home page and an overview page of all clients. The home page (see Figure 4(a)) shows the clients that need special attention. In this case, these are the users that have not logged in onto the platform for a while or have not entered information about the detected stress events. The user should continue to use the platform in agreement with the professional, even when their mental state is improving. The overview page (see Figure 4(b)) shows the information of all clients. For each user, the amount of entered stress events that were annotated relative to the total registered amount is given, the last time that they logged in onto the platform, and β of the HRV (refer to Section 2.4, 2.a). Pressing on the gear icon gives the mental healthcare professional the possibility to change some settings and information about the client (see Figure 4(a)). Here, the step goal per day and per hour can be altered. Also, the amount of detected stress events shown per day can be configured, and this way, only the most severe ones are shown to not overburden the client. Professionals can also extract PDF files that give an overview of the different elements of the platform.

Currently, the measurement data has to be uploaded to the Carewear platform manually. The measurement data is collected on the wearable and afterwards manually uploaded to the Empatica secure cloud platform. There is no rest interface available to connect to this secure platform, so the user has to download the data to his computer manually and subsequently upload it to the Carewear platform. This is typically executed once a day at which time the user also takes the time to go through the reported data and adds its annotations. As mentioned above, only the most severe stress detections of that day are shown, so we need this longer time-frame for defining the most severe stress detections for that day. Asking the user to annotate his feelings shortly after each stress event could present a burden



(a)

Detail view of the day



(b)

FIGURE 2: Continued.

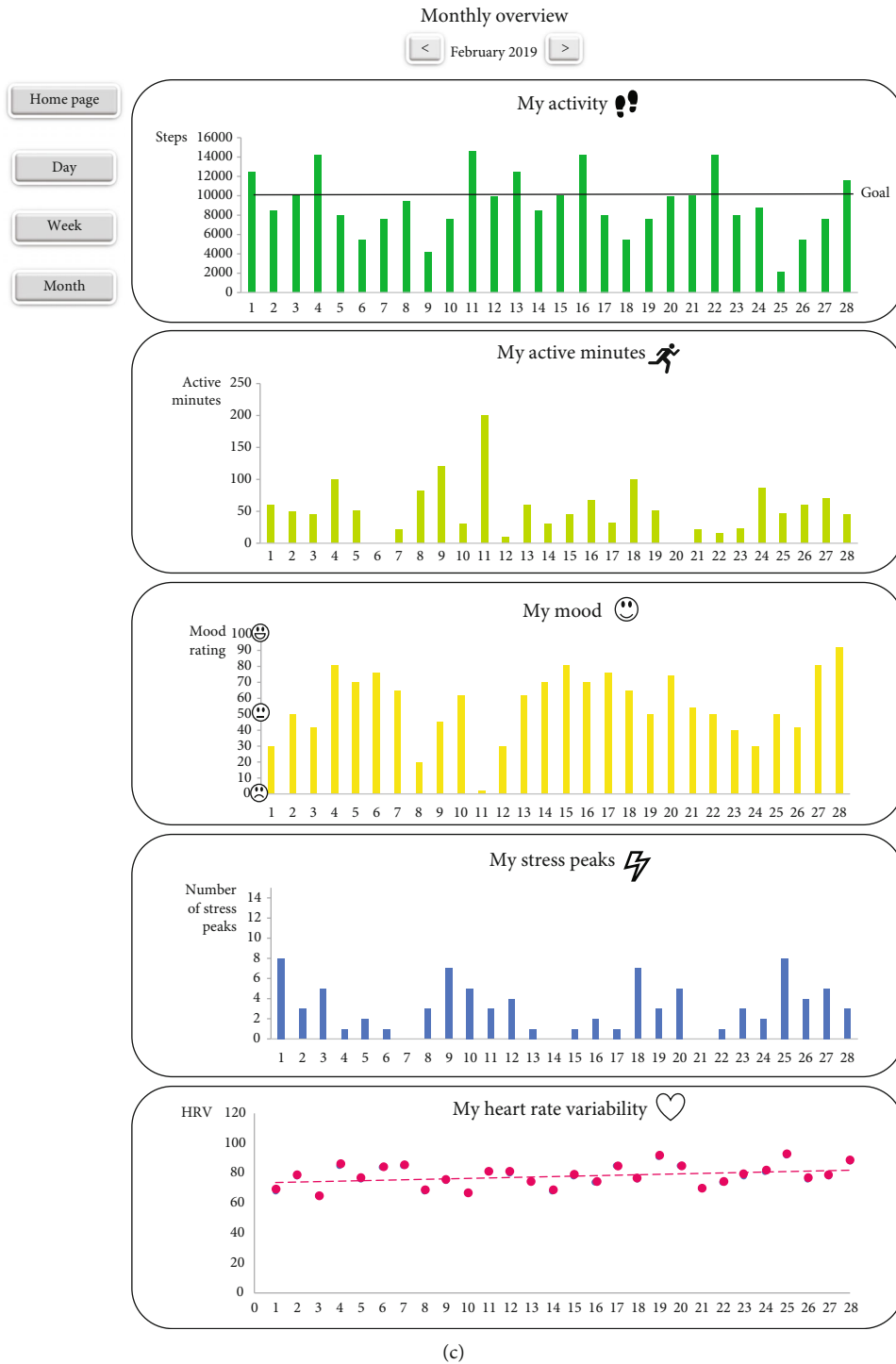


FIGURE 2: Screen caps of the prototype for the Carewear online platform, with the (a) home page, (b) detailed view of the day, and (c) part of the monthly overview.

during stressful times and require highly demanding online computations on the data.

2.4. Algorithms. As stated above, the Empatica E4 measures EDA, BVP, and movement. These are used to calculate the amount of steps, minutes of physical activity,

mean HR, and HRV and to detect moments of acute stress. Before we can extract the different features, the raw signals need to be converted and filtered to reduce the effect of artefacts (and certainly movement artefacts), which is the main challenge in the preprocessing step of the data.

Thursday 7/2 at 9:37 am

Do you remember this moment?
 Yes No

Title:

Did you experience unpleasant stressful feelings?
 Yes No

How did it feel?
 😞 😐 😊

How long did it last?
 h. min.

Notes

Save

FIGURE 3: Screenshot of the prototype for the Carewear online platform on how to enter information for a detected moment of acute stress.

(1) Preprocessing:

- (a) *Accelerometer*: the used accelerometer uses 3 axes. To calculate the strength of movement, the different axes of the 3D-accelerometer signal were combined into one value using the magnitude of the resulting vector calculated as follows:

$$ACC_{MAG} = \sqrt{(axis_x)^2 + (axis_y)^2 + (axis_z)^2} \quad (1)$$

- (b) *EDA*: the EDA is recorded by measuring the skin conductivity. The main challenge with using a wristband is that it is prone to movement artefacts. Not all of these can be removed easily, but to reduce their impact, a low-pass filter is used. The slowly changing part of the EDA signal is called the skin conductance level (SCL) and is a measure of psychophysiological activation. A fast change in the EDA signal (a “peak”) occurs in reaction to a single stimulus (e.g., a startle event) and is called (specific) skin conductance response (SCR). It appears between 1.5 and 6.5 s after the stimulus [45]. Given that we are interested in the moments of acute stress that causes these peak responses in this study, the slowly changing part of the signal is removed. This is done using a high-pass filter. Both filters are combined in a second order Butterworth band-pass filter with

a lower cutoff frequency of 0.05 Hz and a higher cutoff frequency of 5 Hz as used by Kyriakou et al. [8] and Setz et al. [45].

- (c) *BVP to HR*: the BVP is measured using a PPG sensor. This still needs to be converted to HR. Empatica also provides a conversion of the BVP to HR, but we use our own algorithms as their algorithm uses a sliding window to filter the signal that is too large for our purpose. Again, movement artefacts have to be suppressed as good as possible. A second order Butterworth band-pass filter with cutoff frequencies that correspond to a minimal HR of 15 beats per minute (bpm) and a maximal HR of 240 bpm is used for this purpose. On the filtered signal, a peak detection algorithm is used to detect the peaks in the BVP. The time between these peaks is called the interbeat interval (IBI). The HR is the inverse of the IBI, so it can be calculated as such. A failure to detect a peak or an erroneously detected additional peak has a large impact on the detected HR, so the IBIs that produce an HR that are outside of the range of 15 till 240 bpm are removed.
- (2) *Features*: using these processed signals, we calculate the mean HR, the HRV and the change in HRV over time, step count and amount of physical activity, and finally the detected moments of acute stress.

- (a) *HRV and mean HR*: as the detection of the BVP peaks is prone to movement artefacts, we ask the person to sit down and remain as still as possible for a period of 10 minutes each day to measure resting HRV. To label these periods, the user has to press the button on the wearable twice at the start and at the end of the period. The mean resting HR is also calculated over this 10-minute period. The HRV is defined as the root mean square of successive differences (RMSSD), so it can be calculated over the whole period as follows:

$$HRV = \sqrt{\frac{\sum_{n=0}^t (IBI_{n+1} - IBI_n)^2}{n}} \quad (2)$$

Since previous research [36, 37] has shown that a low HRV can be related to stress and mental illness, reductions in HRV could be indicative of decreasing mental health while increases in HRV could be indicative of increased mental health or recovery. However, such interpretation of changes in HRV over time still needs scientific validation. To show the information of an ongoing increase or decrease to the mental healthcare professional, we also calculate a value β which is defined as the slope of the linear regression line

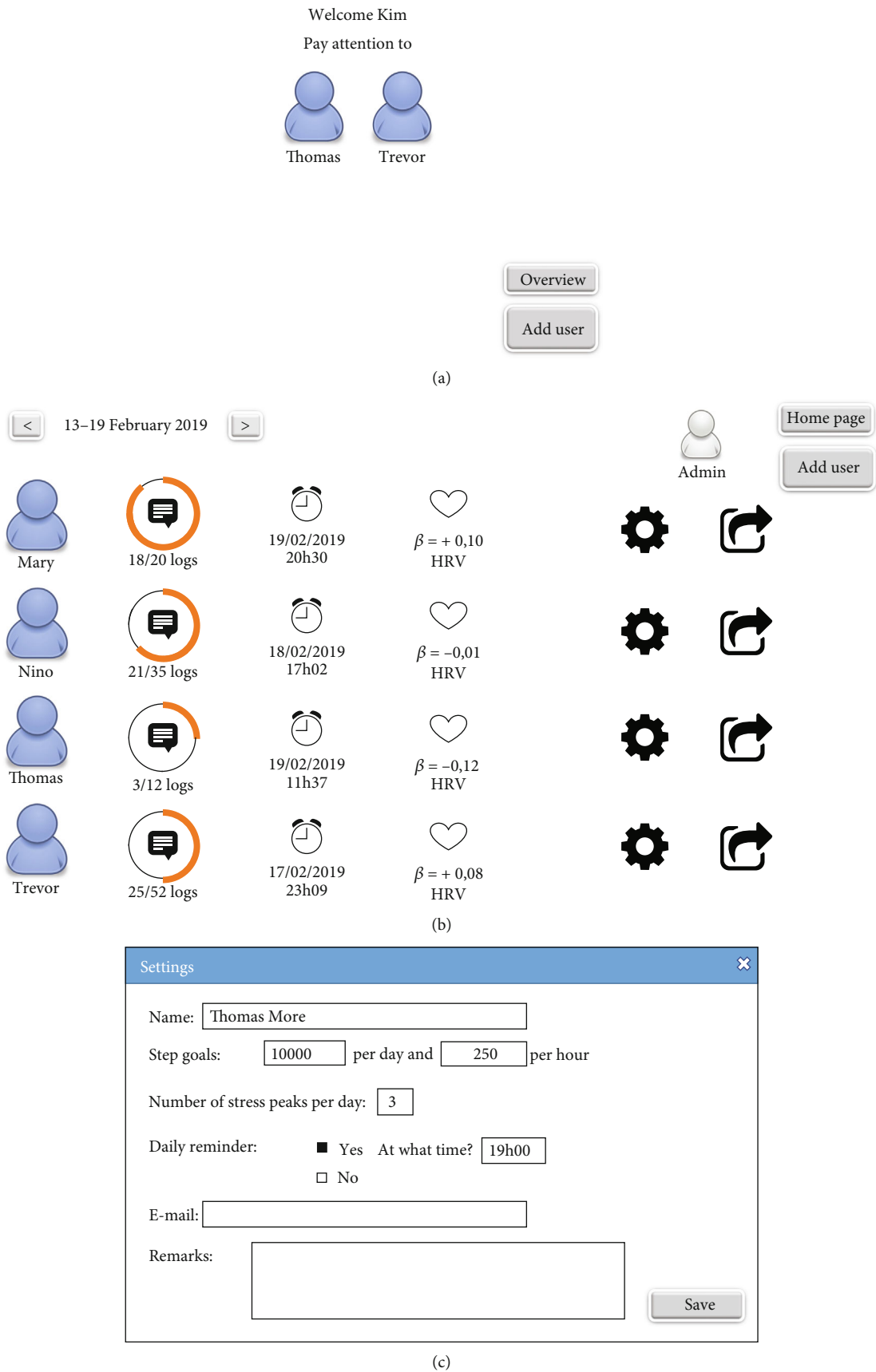


FIGURE 4: Screenshots of the prototype for the Carewear online platform, with the (a) home page of the professional, (b) his detailed view of his users, and (c) the setting configuration of a user.

over all measured HRV values of the monitored person.

- (b) *Step count+physical activity*: physical activity is not only an important factor for the healing process of a client, but also for prevention and overall well-being. The amount of steps a person has walked per hour and per day are shown. Also, the periods of increased physical activity are shown to the client and their mental healthcare professional.

For measuring the amount of steps, a simple peak detection algorithm using ACC_{MAG} was implemented. Every acceleration peak that is greater than 1.375 times the earth gravity is detected as step. This threshold has been determined by tests executed by several persons during multiple periods of 8 hours with several Fitbit devices used as ground truth. This relatively simple approach was used since the current application is mainly interested in trends and relative changes as opposed to the absolute value. Previous research has shown that the absolute step also differs between (commercial) wearable devices and the gold standard [52].

For detecting physical activity, indicative of doing sports such as walking, running, or riding a bicycle, we started by using the approach as documented by Rahman et al. [17]. They defined a threshold on the acceleration energy to detect physical activity. If the standard deviation of the energy of ACC_{MAG} is greater than 0.21348, this is labeled as nonstationary (i.e., walking or running) and the others are labeled stationary. During the tests with different Fitbit devices used as ground truth, this seemed too high for the accelerometer integrated in the Empatica E4. We found that this threshold multiplied by 0.835 gave results that were more consistent with those reported by the Fitbit devices. All periods, in which the mean of this standard deviation is higher than the threshold for longer than nine minutes, are reported as active periods.

- (c) *Moments of acute stress*: for detecting moments of acute stress, we use a combination of SC, HR and, movement. As mentioned above, our body reacts with an increase in SC after a stimulus that invokes arousal. To capture this, we use the first-order derivative $f^t(SC)$ of this signal. Also, the HR reacts in the same way; therefore, the derived signal $f^t(HR)$ is also used here. For the amount of movement, we are only interested to determine if this peak is related to stress or could be related to physical activity. Therefore, we only need the magnitude of the signal, so ACC_{MAG} is used as is.

TABLE 1: Parameters for probability density functions (pdf) of different signals used for detecting moments of acute stress.

Signal	μ	σ
$f^t(SC)$	$\ln(0.35)$	0.6
$f^t(HR)$	10	2.5
ACC_{MAG}	0.981	0.3

For the development of the stress probability algorithm, sixteen healthy individuals were exposed to two different stress inductions, consisting of the Montreal Imaging Stress Task [53] and anxiety-inducing VR clips. Note that nobody of these sixteen individuals was included in the current study. The participants (who provided informed consent) and the observing researcher each noted instances and indications of stress. The physiological signals were analyzed and annotated manually to label the occurring stress peaks during these instances. Using these stress induction experiments, we determined for each signal a probability density function (PDF). This PDF shows the chance that this signal value corresponds to an actual moment of acute stress. It is modeled using a Gaussian distribution with a given mean value μ and standard deviation σ shown in Table 1.

To detect the possible moments of acute stress, a peak detection algorithm is used on $f^t(SC)$. Bursts of peaks were grouped using nonmaximum suppression to only use the strongest peak. After this, for each detected peak, the probability that this peak is related to a stress response is determined using a combination of the three probabilities. For the HR, the maximum probability in a window of 60 seconds before the SC peak is chosen. For ACC_{MAG} , the minimum probability in a window of ten seconds before the peak is used. This way, peaks that could be caused by movement and movement artefacts receive a lower probability. This ten-second window is based on the approach of Rahman et al. [17]. The probability that this detected peak is a stress peak is then calculated as follows:

$$P(SP) = P(SC) \cdot \min(P(ACC_{MAG})) \cdot \max(P(HR)). \quad (3)$$

Each detected peak is thus given a certain probability. To not overload the user, only the three moments of acute stress with the highest probability of each day are shown to the Carewear application user. However, the number of presented moments of acute stress can be tailored to the user by the professional in the platform, if needed.

TABLE 2: Results.

Participant	TP	FP	FN
1	5	14	1
2	2	10	0
4	6	11	1
5	4	16	4
6	6	11	4
Total	23	62	10

TABLE 3: Results for the optimal threshold ($P(SP) > 0.5$).

Participant	TP	FP	FN
1	5	14	1
2	2	10	0
4	6	10	1
5	4	8	4
6	6	11	4
Total	23	53	10

3. Results

First, the results from the five participants of this study are shown; afterwards, the feedback received from them is elaborated on.

3.1. Results of the Study with Five Healthy Participants.

Table 2 gives an overview of the results from our small sample. A total of 23 stress events or true positives (TPs) are detected correctly. Some stress events that were manually logged by the user are not detected giving 10 false negatives (FNs). Part of these false negatives is because only the, in this case, three detections with the highest probability were shown to the user. Also, 62 events are erroneously detected as false positive (FP) stress events. This gives a recall or sensitivity of 0.7 and a precision or positive predictive value (PPV) of 0.27.

As stated before, a total of three moments of acute stress per day are returned. This can introduce a number of false positives. For the moment, we do not take into account how high the probability of a detected moment of acute stress is, and the three highest ones are shown irrespective of their probability. Participant 5 for example has two days in which the highest probability is only 0.32. Only those showing detections higher than a certain threshold can reduce the amount of false positives. Using a threshold of 0.5 gives the best trade-off between precision and recall in this study. The results are shown in Table 3. This gives a recall of 0.7 and a precision of 0.30.

3.2. *Feedback from Users.* The participants were asked for their feedback to improve the platform and user experience. Most participants found the Empatica E4 wearable quite big. Another drawback was that it does not provide direct feedback to the user. The electrodes for the EDA sensor are

large, and they press into the skin, which can get painful after wearing it for a prolonged period of time. This is especially so, given the fact that the wearable has to be worn tight to get measurements of a sufficient quality with less artefacts. This furthermore causes increased sweating under the wristband, which increases the uncomfortable feeling. We provided a manual on how to install the Empatica software on the user's computer, but in a few cases, an error was returned during the installation. To solve these errors, an intervention from a more technically oriented person was needed. Another drawback is that transferring the data from the wearable to the Carewear platform is quite time-consuming and needs manual input from the user. This process can take up to twenty minutes for the complete transfer. Once the data is uploaded to the Carewear platform, it still needs to be processed by the algorithms before being visible to the user. The processing time for this depends on the amount of data and on the data itself. For the HRV measurement, currently the user has to double tag the Empatica E4 button at the start and end of the segment; however, this was not always successfully executed. The users also had some doubts about some of the events that were detected and some not, but these are related to the false positives and negatives as shown above. They however found that the distinction between stress events and physical activity worked rather well. Annotating the detected moments of acute stress was sometimes difficult, because the users could not remember what exactly happened at that particular time. Taken together, usability can be improved by having a more comfortable wearable device as well as having easier (and faster) data transfer and processing procedures.

4. Discussion

This study is limited in the number of participants, which is, however, not uncommon. A recent review by Larradet et al. [4] showed that studies focusing on stress in real life have a mean of 13 participants with a standard deviation of 10. Also, all of them were recruited from the general population and indicated that they did not experience a lot of stress while wearing the device. This contributed to the substantial amount of false positives and explains why they sometimes had difficulty remembering what exactly happened at the detected moments of acute stress. The validity assessment of van Lier et al. [54] also states that the Empatica E4 works best for large stressors, so future research should ascertain the performance of the stress algorithm in a large group of participants including persons with stress-related complaints.

Determining the exact causes of a FP or FN is quite difficult, but in some cases, it is still feasible. A considerable part of the stress events that were manually added by the participants but not detected as a stress event happened while moving, for example, being nervous to catch a train while running to the platform. These kinds of stress events are difficult to detect, as physical activity introduces noise on the signal and can confound the assessment of stress [17]. Can et al. also state that the quality of the BVP signal declines drastically in the case of intense physical activities [44]. This is also

the reason why we give these segments a lower probability in our algorithm. The false detections were mainly caused by noise on the measured signals, certainly the SC and HR. Another part of the FP was during taking classes, so it is possible that these are cognitive tasks that are detected as stress events. As shown in Table 3, the amount of FP can partially be reduced by only showing the detections that have a probability that is higher than a certain threshold. The threshold has to be kept low in this healthy sample, but in clients with stress-related complaints, this could be increased while still detecting the most stressful events. Also, some cases of invalid data were seen, which was caused by not correctly fixating the wearable. Participant 2, for example, was asked to use the wearable again for a week, as most data was unreliable the first time because of this reason. However, given that the study consisted of naturalistic data collection (without stress induction or deception) and that removing this participant from the study results only presented minor changes with a recall of 0.68 and a precision of 0.29, this individual was included in the data analysis.

We collected two types of self-report information from the participants. Firstly, we asked a continuous registration of their activities, and secondly, the platform offered them automated stress detections that they needed to annotate. This constant monitoring and awareness of the data could potentially confound the naturalistic experience. The reason why we did not choose for random sampling in and experience sampling approach is that it was paramount to gain information on their activities during the stress detections to assess the performance of the algorithms. Given that data collection only occurred during one week, random sampling would have entailed substantial loss of data.

A limitation is that we did not make use of the skin temperature in our algorithms, which can, according to Sano et al. [43], aid in classifying between high- and low-stress groups. Using additional information about the ST could enhance the performance of our detection algorithms and can add additional parameters from the different captured signals. However, for the moment the algorithm is manually implemented using rules and probabilities, adding additional parameters and sufficiently tuning these rules and probabilities are not easy. After obtaining more (labeled) data, we could use machine learning techniques to better cope with additional information and further increase the performance. Can et al. [3] state that models that are trained using data collected in lab conditions outperform models that are solely based on real-life data. Thus, additional data should be collected in the lab and in real life. As also stated by Larradet et al. [4], using a continuous monitoring system in real life allows for iterative and personalized learning. Using these data can not only be used to improve the global model, but also to personalize the model to take into account the exact parameters (e.g., age, sex, fitness, etc.) of the monitored person.

As mentioned above, the usability of the platform needs some improvement. For example, the double tagging for the HRV measurements should be changed to an automatic detection of resting periods. A current major drawback in the usage of our platform is the fact that the user has to exe-

cute a chain of manual transactions to transfer the data to the Carewear platform. However, there is currently no REST API available from Empatica, so automatically retrieving the data from the Empatica platform is not possible. A solution for this would be the development of an application that runs on a smartphone which has a continuous connection with the Empatica E4 using a BLE connection. The drawback of this is that the battery of both the smartphone and the Empatica E4 will be depleted sooner.

It is important to continue monitoring the market for devices that provide the same set of sensors in a more easy-to-wear package. For example, Empatica is currently working on a new wearable, the EmbracePlus [55], which could potentially solve some of the challenges we experienced with the Empatica E4. Also, Fitbit introduces the Fitbit Sense with additional sensors [56]. As mentioned before, we believe that stress detection should be done continuously in real life and on a broad scale, so it is important to use commercially available devices that are readily available and easy to use and wear. Gradl et al. [40] rated wearables that are able to measure EDA highest. But they believe that cheaper and more commercial devices, such as basic Fitbit trackers or Apple smartwatches, also have potential to be used for stress detection.

A final limitation is that stress detections and related contextual information do not inform us about how an individual recovers from this stressor or what the long-term impact might be. Future research should also be aimed at informing on physiological and affective stress recovery in real life.

5. Conclusion

Stress is a complex combination of physiological, behavioral, and emotional responses. And although experiencing stress by itself is not maladaptive, prolonged experience of stress without sufficient recovery can impede daily-life functioning and contribute to the development of mental illness. Capturing all stressful events accurately using physiological responses only is not feasible at this point. Combining detected stress events with self-report is key to improve accuracy. To help individuals cope stressful situations, actionable data is needed through adding context to these stress events. From the first pilots and this study, we can conclude that the current application can collect information about stressful situations and their context, using objective data collected in real life, and allow professionals and their clients to interpret the data, observe patterns, and provide tailored interventions.

While the current study shows that a substantial amount of false positives are detected in a healthy sample and that usability could be improved, the concept of a user-friendly platform to combine physiological data with self-report to inform on stress and mental health is viewed positively.

Data Availability

The aggregated data supporting the conclusions of this manuscript are available upon request to the corresponding

author, without undue reservation, to any qualified researcher.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

This work was funded by a VLAIO TETRA grant (HBC.2016.0099) from Flanders Innovation & Entrepreneurship (VLAIO). Companies and organizations involved in the project were ALMA.care, Aptus, Flemish Organization for Psychological Consultants (Beroepsvereniging voor Psychologisch Consulanten), COMmeto, Faresa, IDEWE, IMEC, Pulso Europe, Learn2Act, SD Worx, Securex, The Human Link, and Flemish Organization of Clinical Psychologists (Vlaamse Vereniging van Klinisch Psychologen).

References

- [1] J. E. Haberer, T. Trabin, and M. Klinkman, "Furthering the reliable and valid measurement of mental health screening, diagnoses, treatment and outcomes through health information technology," *General Hospital Psychiatry*, vol. 35, no. 4, pp. 349–353, 2013.
- [2] T. Van Daele and T. Vanhoomissen, "Portable technology in mental healthcare," *De psycholoog*, vol. 7, pp. 34–39, 2015.
- [3] Y. S. Can, D. Gokay, D. R. Kılıç, D. Ekiz, N. Chalabianloo, and C. Ersoy, "How laboratory experiments can be exploited for monitoring stress in the wild: a bridge between laboratory and daily life," *Sensors*, vol. 20, no. 3, p. 838, 2020.
- [4] F. Larradet, R. Niewiadomski, G. Barresi, D. G. Caldwell, and L. S. Mattos, "Toward emotion recognition from physiological signals in the wild: approaching the methodological issues in real-life data collection," *Frontiers in Psychology*, vol. 11, p. 1111, 2020.
- [5] F.-T. Sun, C. Kuo, H.-T. Cheng, S. Buthpitiya, P. Collins, and M. Griss, "Activity-aware mental stress detection using physiological sensors," in *Mobile Computing, Applications, and Services*, pp. 282–301, Springer, 2012.
- [6] E. Smets, E. Rios Velazquez, G. Schiavone et al., "Large-scale wearable data reveal digital phenotypes for daily-life stress detection," *NPJ Digital Medicine*, vol. 1, no. 1, p. 67, 2018.
- [7] Y. S. Can, B. Arnrich, and C. Ersoy, "Stress detection in daily life scenarios using smart phones and wearable sensors: a survey," *Journal of Biomedical Informatics*, vol. 92, article 103139, 2019.
- [8] K. Kyriakou, B. Resch, G. Sagl et al., "Detecting moments of stress from measurements of wearable physiological sensors," *Sensors*, vol. 19, no. 17, p. 3805, 2019.
- [9] E. J. Urban, S. T. Charles, L. J. Levine, and D. M. Almeida, "Depression history and memory bias for specific daily emotions," *PLoS One*, vol. 13, no. 9, article e0203574, 2018.
- [10] J. Alexander, M. McAllister, and D. L. Brien, "Exploring the diary as a recovery-oriented therapeutic tool," *International Journal of Mental Health Nursing*, vol. 25, no. 1, pp. 19–26, 2016.
- [11] T. Debeuf, S. Verbeken, M.-L. van Beveren, N. Michels, and C. Braet, "Stress and eating behavior: a daily diary study in youngsters," *Frontiers in Psychology*, vol. 9, p. 2657, 2018.
- [12] M. S. Patel, L. Foschini, G. W. Kurtzman et al., "Using wearable devices and smartphones to track physical activity: initial activation, sustained use, and step counts across sociodemographic characteristics in a national sample," *Annals of Internal Medicine*, vol. 167, no. 10, pp. 755–757, 2017.
- [13] E. T. Kersten-van Dijk, J. H. D. M. Westerink, F. Beute, and W. A. IJsselstein, "Personal informatics, self-insight, and behavior change: a critical review of current literature," *Human-Computer Interaction*, vol. 32, no. 5-6, pp. 268–296, 2017.
- [14] Y. Ranjan, Z. Rashid, C. Stewart et al., "Radar-base: open source Mobile Health platform for collecting, Monitoring, and analyzing data using sensors, wearables, and mobile devices," *JMIR mHealth and uHealth, under review*, vol. 7, no. 8, p. e11734, 2019.
- [15] H. J. Han, S. Labbaf, J. L. Borelli, N. Dutt, and A. M. Rahmani, "Objective stress monitoring based on wearable sensors in everyday settings," *Journal of Medical Engineering & Technology*, vol. 44, no. 4, pp. 177–189, 2020.
- [16] K. Hovsepian, M. al'Absi, E. Ertin, T. Kamarck, M. Nakajima, and S. Kumar, "cStress: towards a gold standard for continuous stress assessment in the mobile environment," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*, pp. 493–504, Zurich, Switzerland, 2015.
- [17] M. M. Rahman, D. H. Epstein, K. L. Preston et al., "Are we there yet? Feasibility of continuous stress assessment via wireless physiological sensors," in *Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics - BCB '14*, pp. 479–488, Newport Beach California, USA, 2014.
- [18] Y. Tazawa, K.-c. Liang, M. Yoshimura et al., "Evaluating depression with multimodal wristband-type wearable device: screening and assessing patient severity utilizing machine-learning," *Heliyon*, vol. 6, no. 2, article e03274, 2020.
- [19] F. de Arriba-Pérez, J. M. Santos-Gago, M. Caeiro-Rodríguez, and M. Ramos-Merino, "Study of stress detection and proposal of stress-related features using commercial-off-the-shelf wrist wearables," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 12, pp. 4925–4945, 2019.
- [20] V. Mishra, G. Pope, S. Lord et al., "Continuous detection of physiological stress with commodity hardware," *ACM Transactions on Computing for Healthcare*, vol. 1, no. 2, pp. 1–30, 2020.
- [21] A. Muaremi, B. Arnrich, and G. Tröster, "Towards measuring stress with smartphones and wearable devices during workday and sleep," *BioNanoScience*, vol. 3, no. 2, pp. 172–183, 2013.
- [22] E. Smets, P. Casale, U. Großekathöfer et al., "Comparison of machine learning techniques for psychophysiological stress detection," in *Pervasive Computing Paradigms for Mental Health. MindCare 2015*, pp. 13–22, Springer, 2015.
- [23] F. H. Wilhelm, M. C. Pfaltz, P. Grossman, and W. T. Roth, "Distinguishing emotional from physical activation in ambulatory psychophysiological monitoring," *Biomedical Sciences Instrumentation*, vol. 42, pp. 458–463, 2006.
- [24] R. Wang, M. S. H. Aung, S. Abdullah et al., "CrossCheck: toward passive sensing and detection of mental health changes in people with schizophrenia," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 886–897, Heidelberg, Germany, 2016.
- [25] J. T. O'Brien, P. Gallagher, D. Stow et al., "A study of wrist-worn activity measurement as a potential real-world

- biomarker for late-life depression,” *Psychological Medicine*, vol. 47, no. 1, pp. 93–102, 2017.
- [26] M. R. Song, Y.-S. Lee, J.-D. Baek, and M. Miller, “Physical activity status in adults with depression in the National Health and Nutrition Examination Survey, 2005–2006,” *Public Health Nursing*, vol. 29, no. 3, pp. 208–217, 2012.
- [27] V. Dibia, “Foqus: a smartwatch application for individuals with adhd and mental health challenges,” in *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 311–312, Reno Nevada, USA, 2016.
- [28] A. Bali and A. S. Jaggi, “Clinical experimental stress studies: methods and assessment,” *Reviews in the Neurosciences*, vol. 26, no. 5, pp. 555–579, 2015.
- [29] G. P. Chrousos and P. W. Gold, “The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis,” *JAMA*, vol. 267, no. 9, pp. 1244–1252, 1992.
- [30] M. N. Jarczok, M. Jarczok, D. Mauss et al., “Autonomic nervous system activity and workplace stressors—a systematic review,” *Neuroscience and Biobehavioral Reviews*, vol. 37, no. 8, pp. 1810–1823, 2013.
- [31] H.-P. Lin, H.-Y. Lin, W.-L. Lin, and A. C.-W. Huang, “Effects of stress, depression, and their interaction on heart rate, skin conductance, finger temperature, and respiratory rate: sympathetic-parasympathetic hypothesis of stress and depression,” *Journal of Clinical Psychology*, vol. 67, no. 10, pp. 1080–1091, 2011.
- [32] T. Föhr, J. Pietilä, E. Helander et al., “Physical activity, body mass index and heart rate variability-based stress and recovery in 16 275 Finnish employees: a cross-sectional study,” *BMC Public Health*, vol. 16, no. 1, p. 701, 2016.
- [33] T. Josefsson, M. Lindwall, and T. Archer, “Physical exercise intervention in depressive disorders: meta-analysis and systematic review,” *Scandinavian Journal of Medicine & Science in Sports*, vol. 24, no. 2, pp. 259–272, 2014.
- [34] T. Teisala, S. Mutikainen, A. Tolvanen et al., “Associations of physical activity, fitness, and body composition with heart rate variability-based indicators of stress and recovery on workdays: a cross-sectional study,” *Journal of Occupational Medicine and Toxicology*, vol. 9, no. 1, p. 16, 2014.
- [35] M. Lindwall, M. Gerber, I. H. Jonsdottir, M. Börjesson, and G. Ahlborg, “The relationships of change in physical activity with change in depression, anxiety, and burnout: a longitudinal study of swedish healthcare workers,” *Health Psychology*, vol. 33, no. 11, pp. 1309–1318, 2014.
- [36] J. F. Thayer, F. Åhs, M. Fredrikson, J. J. Sollers III, and T. D. Wager, “A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health,” *Neuroscience & Biobehavioral Reviews*, vol. 36, no. 2, pp. 747–756, 2012.
- [37] L. Tonello, F. B. Rodrigues, J. W. S. Souza, C. S. G. Campbell, A. S. Leicht, and D. A. Boulosa, “The role of physical activity and heart rate variability for the control of work related stress,” *Frontiers in Physiology*, vol. 5, p. 67, 2014.
- [38] A. H. Kemp, D. S. Quintana, M. A. Gray, K. L. Felmingham, K. Brown, and J. M. Gatt, “Impact of depression and antidepressant treatment on heart rate variability: a review and meta-analysis,” *Biological Psychiatry*, vol. 67, no. 11, pp. 1067–1074, 2010.
- [39] A.-K. Lennartsson, I. Jonsdottir, and A. Sjörs, “Low heart rate variability in patients with clinical burnout,” *International Journal of Psychophysiology*, vol. 110, pp. 171–178, 2016.
- [40] S. Gradl, M. Wirth, R. Richer, N. Rohleder, and B. M. Eskofier, “An overview of the feasibility of permanent, real-time, unobtrusive stress measurement with current wearables,” in *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, pp. 360–365, Trento, Italy, 2019.
- [41] J. Wijsman, B. Grundlehner, H. Liu, H. Hermens, and J. Penders, “Towards mental stress detection using wearable physiological sensors,” in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1798–1801, Boston, MA, USA, 2011.
- [42] Q. Xu, T. L. Nwe, and C. Guan, “Cluster-based analysis for personalized stress evaluation using physiological signals,” *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 1, pp. 275–281, 2015.
- [43] A. Sano, S. Taylor, A. W. McHill et al., “Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: observational study,” *Journal of Medical Internet Research*, vol. 20, no. 6, article e210, 2018.
- [44] Y. S. Can, N. Chalabianloo, D. Ekiz, and C. Ersoy, “Continuous stress detection using wearable sensors in real life: algorithmic programming contest case study,” *Sensors*, vol. 19, no. 8, p. 1849, 2019.
- [45] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Tröster, and U. Ehlert, “Discriminating stress from cognitive load using a wearable EDA device,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 410–417, 2010.
- [46] M. Gjoreski, H. Gjoreski, M. Lus̆trek, and M. Gams, “Continuous stress detection using a wrist device: in laboratory and real life,” in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pp. 1185–1193, Heidelberg, Germany, 2016.
- [47] E. Kaczor, S. Carreiro, J. Stapp, B. Chapman, and P. Indic, “Objective measurement of physician stress in the emergency department using a wearable sensor,” in *Proceedings of the 53rd Hawaii International Conference on System Sciences*, p. 3729, Hawaii, USA, 2020.
- [48] N. A. J. De Witte, “Stress en depressie te lijf met wearables. handleiding om evidence-based aan de slag te gaan met carewear [working on stress and depression with wearables. manual for evidence-based use of carewear],” *Antwerpen, Applied Psychology, Thomas More University of Applied Sciences*, 2018.
- [49] N. A. J. De Witte, “Gebruik van de wearable en het carewear-platform. praktische handleiding [manual for the use of the wearable and the carewear platform],” *Antwerpen, Applied Psychology, Thomas More University of Applied Sciences*, 2018.
- [50] Empatica, “Empatica E4 Home page,” March 2020, <https://www.empatica.com/research/e4/>.
- [51] L. Menghini, E. Gianfranchi, N. Cellini, E. Patron, M. Tagliabue, and M. Sarlo, “Stressing the accuracy: wrist-worn wearable sensor validation over different conditions,” *Psychophysiology*, vol. 56, no. 11, article e13441, 2019.
- [52] T. Fokkema, T. J. M. Kooiman, W. P. Krijnen, C. P. Van der Schans, and M. De Groot, “Reliability and validity of ten consumer activity trackers depend on walking speed,” *Medicine & Science in Sports & Exercise*, vol. 49, no. 4, pp. 793–800, 2017.
- [53] K. Dedovic, R. Renwick, N. K. Mahani, V. Engert, S. J. Lupien, and J. C. Pruessner, “The montreal imaging stress task: using functional imaging to investigate the effects of perceiving and

processing psychosocial stress in the human brain,” *Journal of Psychiatry and Neuroscience*, vol. 30, no. 5, pp. 319–325, 2005.

- [54] H. G. van Lier, M. E. Pieterse, A. Garde et al., “A standardized validity assessment protocol for physiological signals from wearable technology: methodological underpinnings and an application to the e4 biosensor,” *Behavior Research Methods*, vol. 52, no. 2, pp. 607–629, 2020.
- [55] Empatica, “Empatica Embrace Plus Home page,” March 2020, <https://www.empatica.com/embraceplus>.
- [56] Fitbit, “Fitbit Sense Home page,” September 2020, <https://www.fitbit.com/be/sense>.