ORIGINAL ARTICLE

CrossMark

Objective detection of chronic stress using physiological parameters

Rabah M. Al abdi[1] (iD) · Ahmad E. Alhitary[2] · Enas W. Abdul Hay[1] · Areen K. Al-bashir[1]

Abstract

The aim of this study was to design a system to diagnose chronic stress, based on blunted reactivity of the autonomic nervous system (ANS) to cognitive load (CL). The system concurrently measures CL-induced variations in pupil diameter (PD), heart rate (HR), pulse wave amplitude (PWA), galvanic skin response (GSR), and breathing rate (BR). Measurements were recorded from 58 volunteers whose stress level was identified using the State-Trait Anxiety Inventory. Number-multiplication questions were used as CLs. HR, PWA, GSR, and PD were significantly ($p < 0.05$) changed during CL. CL-induced changes in PWA (16.87 ± 21.39), GSR (-13.71 ± 7.86), and PD (11.56 ± 9.85) for non-stressed subjects ($n = 36$) were significantly different ($p < 0.05$) from those in PWA (2.92 ± 12.89), GSR (-6.87 ± 9.54), and PD (4.51 ± 10.94) for stressed subjects ($n = 22$). ROC analysis for PWA, GSR, and PD illustrated their usefulness to identify stressed subjects. By inputting all features to different classification algorithms, up to 91.7% of sensitivity and 89.7% of accuracy to identify stressed subjects were achieved using 10-fold cross-validation. This study was the first to document blunted CL-induced changes in PWA, GSR, and PD in stressed subjects, compared to those in non-stressed subjects. Preliminary results demonstrated the ability of our system to objectively detect chronic stress with good accuracy, suggesting the potential for monitoring stress to prevent dangerous stress-related diseases.

Abbreviations

ANS	Autonomic nervous system
BR	Breathing rate
CL	Cognitive load
ECG	Electrocardiogram
EMG	Electromyogram
GSR	Galvanic skin response
HR	Heart rate
HRV	Heart rate variability
PD	Pupil diameter
PPG	Photoplethysmography
PWA	Pulse wave amplitude
SNS	Sympathetic nervous system
STAI	State Trait Anxiety Inventory

✉ Rabah M. Al abdi
  rmalabdi@just.edu.jo

  Ahmad E. Alhitary
  ahmad2005131@gmail.com

  Enas W. Abdul Hay
  ewabdulhay@just.edu.jo

  Areen K. Al-bashir
  akbashir@just.edu.jo

[1] Biomedical Engineering Department, Faculty of Engineering, Jordan University of Science and Technology, Irbid 22110, Jordan

[2] Al-Amal Arabia Specialized Hospital, Al-Hudaydah, Yemen

1 Introduction

Stress generally refers to unpleasant feelings aroused by mental, emotional, or physical challenges that we encounter in our lives [38]. Short-term stress is healthful, since it helps us solve problems that we face in our lives by causing us to stay focused, energetic, and alert. However, chronic stress, associated with a state of anxiety, causes lack of coordination between the human body and the mind and reduces productivity. If chronic stress is not identified and controlled in its early stages, it can cause hypertension [42], increased likelihood of infections [38], depression [44], and cardiovascular diseases [20, 22, 30]. Chronic stress can be controlled by relaxation and with suitable

management techniques, including exercise, healthful eating habits, positive attitudes and goals, and social life, for primary prevention of stress-related diseases. The incidence of chronic stress in modern life is increasing for both old and young people. For example, the stress levels in college students have been raised due to adjusting to college life, dealing with unexpected societal event, and fulfilling developmental tasks, which contributed to physiological and psychological problems [5, 39]. But identifying stress accurately without help from a medical expert remains a challenge. Developing a system for monitoring people to objectively identify chronic stress in its early stages is important to control stress before it causes chronic untreated diseases.

Stress stimulates the sympathetic nervous system (SNS), producing a stress response with psychological and physiological symptoms [19]. Psychological symptoms are related to the mind or mental activity without execution of an action [35]. The psychological symptoms are generally evaluated by subjective methods, based on counseling using a questionnaire [24, 40, 52]. The questionnaire asks about how one generally feels, and the subject circles an appropriate number that indicates his/ her degree of feeling [40]. If a subject reads each question with understanding, and describes his/her actual feelings without bias, the output from the questionnaire gives a good measure of stress [17]. A definitive questionnaire for measuring anxiety in adults is the State-Trait Anxiety Inventory (STAI), which is the leading measure of personal anxiety used worldwide [10]. STAI is simple and ideal for evaluating individuals with chronic stress and anxiety. In this study, the score of STAI was used as a ground truth for identifying subjects with chronic stress.

Physiological symptoms result from normal functioning of bodily parts [34]. Even though physiological symptoms are non-voluntary actions, which are difficult or impossible to be seen by an external observer, they can be detected using biomedical sensors. Several objective techniques have been used to investigate evaluating stress by measuring biomarkers and physiological parameters. Measurements of biomarkers, including cortisol and catecholamines in blood or saliva, were used to evaluate chronic stress in several studies [28, 43]. Catecholamines are known to control the initial response to stress, while glucocorticoids take a longer time to be secreted [3]. People with chronic stress had elevated cortisol levels [23]. However, these studies were invasive, costly, used slow methods of analysis, and required special expertise [26], thus limiting the use of these methods for monitoring stress.

As an alternative way to measure biomarkers, researchers have studied several physiological signals and parameters to identify which of these are sensitive to stress. The physiological signals studied were ECG, EEG, EMG, photoplethysmography (PPG), respiration, electrodermal activity, and pupillometry. Changes in these signals were induced by cognitive loads (CL), including simple motor tasks [36], Stroop color-word conflict tests [26], arithmetic and memory tasks [49], and simulated driving tasks [13, 33]. The physiological parameters extracted from these signals were heart rate (HR) and HR variability, from ECG [49]; HR and pulse wave amplitude (PWA), from PPG [26, 53]; mean amplitudes of event related potential (ERP) components, from EEG [11]; galvanic skin response (GSR), from electrodermal activity [32]; mean and standard deviation of EMG [13]; mean of pupil diameter (PD), from pupillometry [21, 53]; and breathing rate (BR), from respiration [13, 47]. The physiological parameters that changed most significantly due to CL, and provided reliable information about stress levels, were HR [49], PWA [26], GSR [54], and PD [33, 36]. In many of these studies, the words "mental stress" or "stress" were used to indicate CL. However, none of the studies compared CL-induced changes in physiological parameters between people who are stressed or non-stressed as determined by a STAI score.

A recent study reported that chronic stress caused degradation in the autonomic nervous system (ANS) reaction to CL [43], based on measurements of blood cortisol and salivary alpha-amylase concentrations. Building on their conclusion, we hypothesized that the reduced ANS reaction to CL in stressed subjects appears also in physiological parameters. Thus, accurate measurements of CL-induced changes in physiological parameters that are innervated by the SNS may permit the identification of stressed people, based on their blunted SNS reaction to CL.

The aim of this research was to investigate the influence of chronic stress on the variation of physiological parameters due to CL. We designed a system to evaluate SNS responses by simultaneously measuring variations in HR, PWA, BR, GSR, and PD between rest and CL. STAI scores were used as a ground truth to classify subjects as stressed (high STAI scores) or non-stressed [27]. CL-induced changes in physiological parameters were compared between stressed and non-stressed people to identify the parameters sensitive to chronic stress. The goal of this study was to diagnose people with chronic stress by measuring their CL-induced changes in physiological parameters.

Four signals were used to investigate physiological parameters: a forehead PPG to monitor HR and pulse wave amplitude (PWA), respiration to measure BR, grove electrodes to measure galvanic skin response (GSR), and pupillometry to measure pupil diameter (PD). The device sensitivity for PD was optimized based on recommendations from previous studies [15, 41]. The system was calibrated using phantoms, and validated using a human study of 58 healthy volunteers, to document the feasibility of detecting differences in CL-induced changes in physiological parameters between stressed and non-stressed subjects.

2 Materials and methods

2.1 System description and validation

In this section, we describe the design and implementation of a chronic stress detection system, which relies on measuring

physiological parameters noninvasively for the purpose of detecting SNS's degradation caused by chronic stress.

Figure 1 shows a schematic diagram of the system components that include four sensors: 5-megapixel Raspberry Pi camera, reflective forehead plethysmography, two-grove silver plating electrodes, and a thermistor. A personal computer was used to control and store data. Novel elements of the proposed system are imaging the left eye under complete darkness and using a fixed camera-pupil distance to eliminate motion artifacts.

A 5-megapixel Raspberry Pi camera module was mounted on a virtual reality headset to image the left eye from a close distance (4 cm) and under only infrared illumination. Since light variations reached the right eye varied PD in both eyes [31], the right eye was also in darkness. The relative location and distance between the camera and the left eye was fixed even during head and body movements, which eliminated motion artifacts and enabled accurate detection of the pupil diameter. One more advantage of fixing distance in the front of both eyes was reducing accommodation reflex, which was spontaneous changes in pupil diameter when focusing on near and far objects [4, 33]. Near objects required foveal focusing and caused reduction in PD in the range of 1–6 mm [4], which was larger than the expected changes in PD due to cognitive loads (< 2 mm). Figure 2 shows a photograph of the headset mounted on a volunteer's head. A
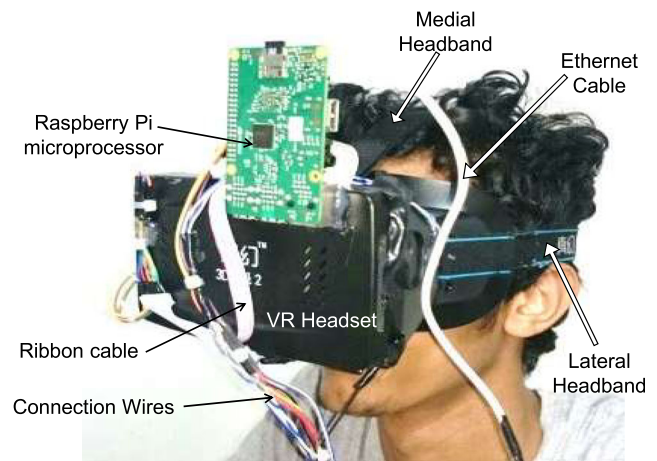

Fig. 2 A photograph of the headset attached to a volunteer's head. VR virtual reality

Raspberry Pi microprocessor board with a 900-MHz quad-core ARM Cortex-A7 CPU and 1 GB RAM was used to capture and buffer images from the camera and send them to the PC via an Ethernet cable with an imaging rate of 15 images per second.

The PPG signal was detected using TCRT1000 reflective optical sensor from Vishay Semiconductors (Tokyo, Japan). Many features of this sensor made it suitable for our measurement and that includes small size, electrically insulated for

Fig. 1 Functional block-diagram of the system. sps samples per second
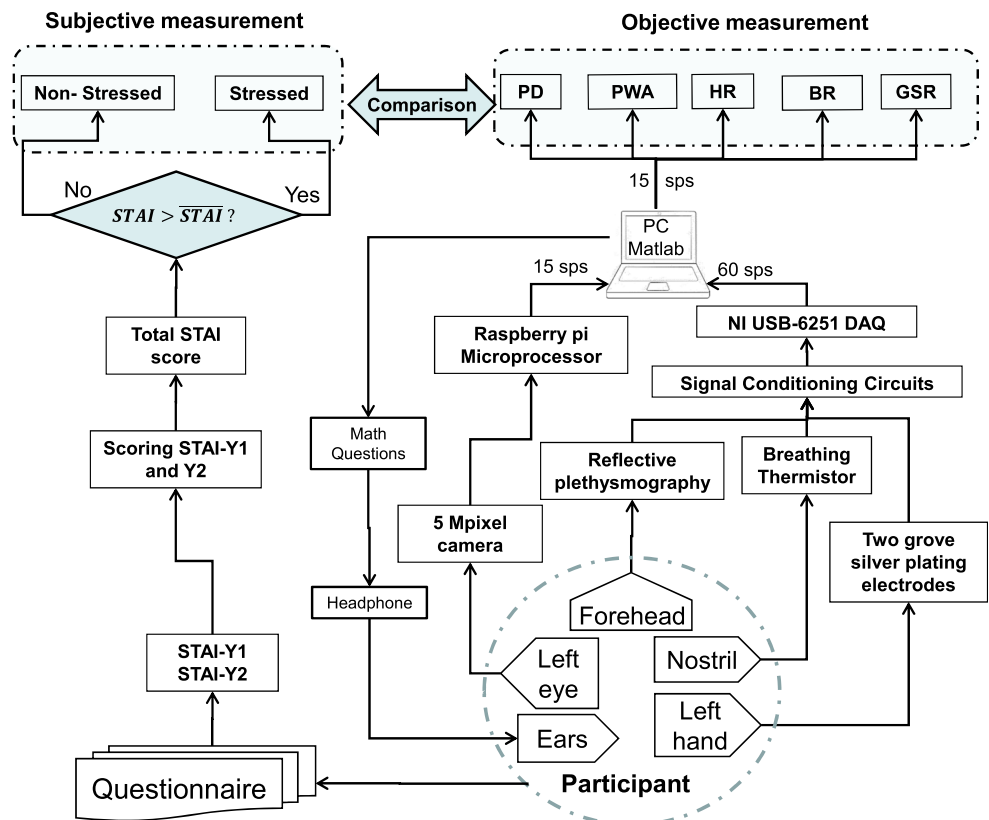
electrical safety, and having a daylight-blocking filter to decrease the interference from room light. The transmitter wavelength is 950 nm, and the receiver type is Phototransistor. An instrumentational amplifier with a common mode rejection ratio of 85 dB was used to reject common mode signals in the PPG signal, and a high-pass filter with a 0.1-Hz cutoff frequency was used to block the large DC component of the PPG signals. Next, a low-pass active filter with a 2.34-Hz cutoff frequency was used to amplify the signal (in the range of − 2 to 2 V) and attenuate power line and high-frequency noise. Then, the PPG signal was converted into digital signals with 16-bit resolution and sent to the PC with 60 samples per second using a data acquisition card USB-6251 DAQ from National Instruments (Austin, TX, USA).

GSR was measured using two grove silver-plated electrodes. A safe electrical current (1 μA) was passed using the two electrodes that were attached to two fingers in the left hand to convert skin resistance into voltage. The voltage across the electrodes was connected into an analog channel in the USB-6251 DAQ to be converted from analog to digital signal with 16-bit resolution and stored on the PC at a sampling frequency of 60 samples per second. The linear operating range was from 0 to 2.1 MΩ.

A small 100-kΩ negative temperature coefficient (NTC) Thermistor from Semitech Semiconductor Ltd. (Melbourne, Australia) was fixed in the headset to monitor respiration. The sensor was placed close to the outlet of the right nostril to detect changes of breath temperature between ambient temperature (inhalation) and lung temperature (exhalation). A voltage divider was first used to convert variations in resistance into voltage by connecting the 100 kΩ NTC thermistor in series with a 100-kΩ carbon film resistance and bias them by 5 V. The voltage across the thermistor increases when the resistance increases (decrease temperature) and decreases when the resistance decreases (increase temperature). This voltage signal that is inversely proportional to the temperature of the inhaled and exhaled air from nostril was inputted into a first-order low-pass active filter with a cutoff frequency of 30 Hz to remove external noise. The output of the amplifier was connected to an analog input of the USB-6251 DAQ to be converted into digital signals with 16-bit resolution and stored on the PC at 60 samples per second.

Knowing that changes in GSR signal tends to decrease very fast due to CL, we sampled the GSR signal with high sampling frequency (60 samples per second) to permit capturing any delay between CL and changes in GSR with high temporal resolution (1/60 s). Since the specification of the USB-6251 DAQ requires the sampling frequencies of all analog input channels to be the equal, the PPG signal was oversampled to keep the sampling frequency the same as GSR signal.

A graphical user interface (GUI) designed in Matlab R2015a software (The MathWorks Inc., MA, USA) was used to control the operation of the system and store data. The system permits 15 high-resolution images per second of the left eye and 60 samples per second with 16-bit resolution of the physiological signals: PPG, breathing, and GSR. During experimental setup, the captured images of the left eye and the physiological signals were being displayed in real time to do any necessary adjustment of the headset and the grove electrodes to ensure the quality of the recorded data.

System calibration was done using laboratory phantoms made from QM Skin 30 (Quantum Silicones Specialties, Richmond, VA). Eyeball-like phantoms (a sphere) were made with 1.6 cm radius as shown in Fig. 3a, and each one had a painted iris and artificial cornea with varying pupil diameter (3 to 8 mm) and iris color. Each phantom was placed 4 cm away from the camera to resemble the actual location of the eye during the study, and 100 images of each phantom were taken after covering the headset and the phantom with a black clothes. The calculated pupil diameter (PD) from the images is modeled as shown in Eq. 1:

$$PD = \varphi + \epsilon \quad \epsilon \sim N(\mu, \sigma) \tag{1}$$

where φ was the true diameter of the pupil (measured directly using a caliper) and ϵ is the measurement error. PD and ϵ were random variables that took on new values for each measurement. The instrument's accuracy was calculated as the absolute value of the bias $|\mu|$, and its stability was calculated as the standard deviation (σ) of the measurement error (ϵ).

Figure 3b shows a photograph of gray slabs of 2 cm thickness and different optical densities that were made from QM Skin 30 silicone; the optical densities were controlled in by adding India ink (changing the silicone color from white to different degree of gray). The darker color of the silicone slab was made to represent darker skin color in human that tended to reflect less light, hence producing smaller PPG signals. The long-term stability of the PPG sensor was calculated by placing a slab over the PPG sensor and getting reading under room light for 15 min. Since there were no changes in the optical properties of each silicone slab, the measured light (reflected light) by the PPG sensor should be constant and does not changes over time. Then, the standard deviation of the PPG signals was calculated to quantify the long-term stability of the PPG readings and their robustness to noise and ambient light.

Stability of the GSR sensor was evaluated by attaching a fixed resistor (0.5 MΩ) between the two GSR electrodes and calculating the standard deviation of a 15-min signal. GSR's accuracy was validated using 18 resistor's value with 0.1 MΩ increasing step of electrical resistance (0.3–2.1 MΩ). The square root of the average square differences between the resistance readings from our device and a Fluke Ohmmeter (Norfolk, UK) was used to evaluate the accuracy of GSR reading. Finally, stability of the thermistor (breathing sensor) was evaluated by immersing the tip of the thermistor in a

Fig. 3 Photographs of the phantoms used for calibrating the camera and the PPG sensor. **a** An eye-like phantom to calibrate the accuracy of measuring pupil diameter. **b** Slap phantoms to evaluate stability of PPG signal

(a)									(b)
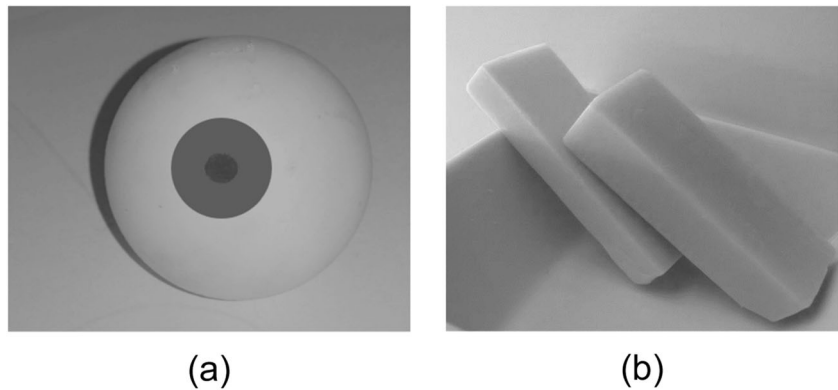
water with fixed and known temperature (23 °C) and calculating the standard deviation of a 15-min signal. On the other hand, accuracy was calculated from the square root of the average square differences between the temperature readings from our thermistor and from a commercial thermometer (Fisher Scientific, USA), over the expected variation in breath's temperature (20–40 °C).

2.2 Participants

Human studies were conducted at Jordan University of Science and Technology (JUST). The study protocol was approved by the Institutional Review Board of JUST. Eligible participants were neither having psychological illness nor drug abuse and were not wearing contact lenses, because they affect PD reading. Fifty-eight healthy students from JUST (26 males, 32 females, 19–25 years old, average 21.52 years) were included in this study. The subjects were asked to refrain from caffeinated food or drinks, smoking cigarettes, and exercising during the 3 hours prior to the experiment.

2.3 Subjective measures of chronic stress

Two forms of STAI for Adult questionnaires (form Y-1 and form Y-2) were used to assess anxiety of subjects included in our study [40]. After gaining the approval from the original author, two health psychologists translated the English versions of STAI-Y1 and STAI-Y2 questionnaires into Arabic language to lessen barriers of assessment with participants. Each form consisted of 20 items to measure anxiety, and a 4-point Likert scale from 1 to 4 was used to rate each item with higher scores indicating high anxiety. These questionnaires were filled by each participant right before the experiment. The total STAI score for each subject was calculated from the sum of both STAI-Y1 and STAI-Y2 scores. Test-retest reliability of the STAI was done by calculating the Pearson correlation coefficient for a subsample of participants ($n = 26$) who repeated the two questionnaires 3 days after the initial administration [2].

We used STAI total score to assess chronic stress level of our subjects. A subject was considered stressed or non-stressed if his/her score was larger or smaller than the average STAI scores of all subjects, respectively [1, 18].

2.4 Experimental setup

All subjects were informed about the nature of the study. Informed consent form and brief medical history were obtained from participants, and the STAI questionnaire was administered. After that, participants were seated in a comfortable adjustable chair positioned sufficiently close to the system to allow for placing the headset on a participant's face in order to image the left eye, read PPG from forehead, and monitor inhalations and exhalations. The two grove electrodes were attached to the middle phalanges of the index and middle fingers of the left hand to measure GSR response. The room of the experiments was quiet, and its temperature was maintained at 23 ± 2 °C.

Adjustable support of the arms and feet was employed to ensure subject comfort and to minimize the tendency of movement. Following this, the operator pressed a START button to begin performing the Math protocol and storing all images and signals on the PC. In addition, the whole experiment was voice-recorded on the PC.

Math protocol was implemented to provoke changes in physiological parameters, by asking subjects to answer number-multiplication questions using headphones. Figure 4 outlines the Math protocol, which includes a baseline of 30 s where subject was sitting quietly and then asking aloud the subject to pronounce the answer of three multiplication questions. To keep consistency, Math questions were played automatically by Matlab in pre-specified time (at 30, 50, and 70 s) from a recorded file. The total time of applying the Math protocol was 90 s, and the difficulty of questions was asked in order: easy (e.g., 6×7), then moderate (e.g., 13×6), and then difficult (e.g., 12×14). The headset of the device was not removed between asking questions. The interval before asking the first question (30 s) was called baseline interval (rest interval), and the time after asking the first question to the end of data collection (60 s) was called cognitive load (CL) interval.

Fig. 4 Timing diagram of the Math protocol implemented for the clinical study. *a.u* Arbitrary unit
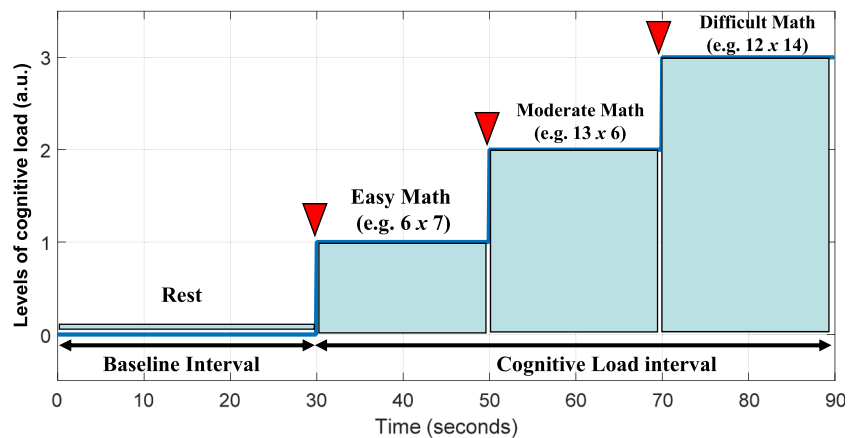
2.5 Data processing and analysis

In this section, we present processing left eye's images and physiological signals to calculate HR, PWA, BR, GSR, and PD and extracting features from them. These processing and feature extraction were done offline using Matlab. In addition, we explain analyzing the influence of CL and stress state on the extracted features using SPSS 19.0 (SPSS, Chicago, IL, USA). Finally, five classification algorithms were implemented to detect chronic stress by inputting all parameters using WEKA software (University of Waikato, New Zealand).

Each image of the left eye was processed to measure PD using the following steps: (1) filtering the image using a novel filtering techniques called 2D total variation denoising (2D-TVD) to filter noise in the image without smoothing sharp edges such as the boundary between pupil area and its surrounding [37, 45], (2) adaptive thresholding to isolate the pupil area and convert the image into binary based on the gray level of each pixel in the image, (3) applying close and fill techniques to remove unwanted scattered regions around the pupil and fill small-undetected points inside the pupil area such as reflected-light spots on the cornea, and (4) detecting pupil edges and finding a best fitted circles for the pupil area to measure PD. It is worth mentioning that eyeblink artifacts encountered during imaging the left eye were detected by a Matlab code based on the average brightness of each image. During eye blinking, the brightness of an image became significantly high as infrared light highly reflected from eyelid. Once an eye blink was identified, the reading of pupil diameter during this blink was calculated using cubic interpolation for the adjacent pupil readings. The number of points in PD time series ($v_{PD}[n]$) was 1350 data points for the whole experiment (15 images per second × 90 s).

Forehead PPG signals were processed to find heart rate (HR) and pulse wave amplitude (PWA). The processing procedures were the following: (1) filtering PPG signals by a fifth-order Butterworth band pass filter with cutoff frequencies of 0.5 and 10 Hz to remove DC wandering, attenuate high-frequency noise, and smooth diacritic notches [48], (2) finding local peaks (minima and maxima) and determine their amplitude and time index, (3) finding HR from the reciprocal of peak-to-peak periods and multiplying the result by 60 to convert HR into beats per minute (bpm), and (4) calculating pulse wave amplitude (PWA) as the difference between the peak (maximum) and nadir (minimum) values of each cardiac cycle [26, 53]. The number of data points in HR and PWA was equal to the number of local peaks in PPG during the experiment (depends on heart rate), and they were between 92 to 168 data points for the whole experiment. Linear interpolation of the readings was used to produce 1350 points for HR time series ($v_{HR}[n]$) and 1350 points for PWA time series ($v_{PWA}[n]$) to match the number of points in $v_{PD}[n]$.

GSR signals were filtered using 1D total variation denoising filter (1D-TVD), which is remarkably effective at simultaneously preserving edges while smoothing away noise in flat regions in GSR signal [37]. There were 5600 data points in GSR time series (60 samples per second × 90 s). Down-sampling of the 5600 points was done to produce a 1350 data points GSR time series ($v_{GSR}[n]$) to match the number of points in $v_{PD}[n]$, $v_{HR}[n]$, and $v_{PWA}[n]$.

Breathing rate (BR) was extracted from the periodic signal coming from the thermistor placed close to nostril. Fifth-order Butterworth low-pass filter with a cutoff frequency of 0.9 Hz was used to smooth the signal and remove unwanted high-frequency noise. Local peaks and time indices of cycles were found using Matlab. BR was calculated from the reciprocal of breath-to-breath intervals. The number of readings in BR ranged from 21 to 37 points, which was dependent on breathing rate. Linear interpolation was used to produce 1350 data points for BR time series ($v_{BR}[n]$).

Five physiological parameters were produced for each subject, and each one contains 1350 data points. To reduce inter subject's variabilities in these physiological parameters, normalization to baseline was implemented. For each time series, normalization was done by subtracting its average value

during baseline interval ($\overline{B}$) from each point in the time series, then dividing the resulted time series by ($\overline{B}$), and after that multiplying the result by 100% to have a percentage changes, as explained in Eqs. 2 and 3:

$$\overline{B} = \frac{\sum_{n=1}^{N1} v_x[n]}{N_1},$$ (2)

$$\acute{v}_x[n] = \frac{v_x[n] - \overline{B}}{\overline{B}} \times 100\%,$$ (3)

where $\acute{v}_x[n]$ denotes the normalized time series, $x = \{$ HR, PWA, BR, GSR, and PD$\}$, and N_1 is the end of baseline interval and was equal to 450 data points (30 s of baseline × 15 points per second). As a result, the normalized time series showed percentage increase or decrease in the physiological parameters relative to their average values during baseline interval. CL-induced changes in the physiological parameters (Δ) were calculated from the average value of $\dot{v}[n]$ during cognitive-load intervals as shown in Eq. 4. These Δs denoted percentage changes in $\acute{v}_x[n]$ between cognitive load and baseline intervals. A cognitive load interval was defined as the time from asking the first Math question ($N_1 = 450$) to the end of the experiment ($N_2 = 1350$):

$$\Delta_x = \frac{\sum_{n=N_1}^{N_2} \dot{v}_x[n]}{N_2 - N_1 + 1}$$ (4)

For each subject, five features were produced: Δ_{HR}, Δ_{PWA}, Δ_{BR}, Δ_{GSR}, and Δ_{PD}. The one-sample T test was used to test if the average across all subjects for Δ was significantly differed from zero, in order to specify features sensitive to cognitive loads. The two-sample T test (two-tailed) was used to determine significant difference in Δ between stressed and non-stressed subjects, in order to specify features that were sensitive to chronic stress. The area under the receiver-operating characteristic curve (AUC) was used to illustrate the diagnostic ability for each individual feature (Δ_{HR}, Δ_{PWA}, Δ_{BR}, Δ_{GSR}, and Δ_{PD}) to discriminate stressed subjects from non-stressed subjects, and the Spearman's rank correlation coefficient RHO was calculated to determine the correlation between Δ_{HR}, Δ_{PWA}, Δ_{BR}, Δ_{GSR}, and Δ_{PD}, and STAI scores. Finally, five classification algorithms were used to combine the diagnostic power of all features in order to discriminate stressed from non-stressed subjects. The used algorithms were logistic linear regression, Naïve Bayes, Multilayer Perceptron, Random forest, and K-star (K^*) [25]. The WEKA software, which can be freely downloaded from http://www.cs.waikato.ac.nz/ml/weka, was utilized for the classification process [9]. The classification performance was evaluated using 10-fold cross-validation method to obtain

a more accurate and realistic assessment of the classifiers [50]. Ten-fold cross-validation was based on randomly partitioning the 58 subjects into 10 subsamples. Eight of the subsamples had six subjects and two of the subsamples had five subjects. Of the ten subsamples, nine were used to train a classifier, and the tenth was used to validate the classifier. This process was repeated 10 times—each of the 10 subsamples was used exactly nine times for training and one time for validating. The 10 results obtained from applying classifiers on the validating subsamples were combined to produce a confusion matrix that contains information about actual and classified states of subjects. We calculated the sensitivity of each classifier by dividing the number of subjects correctly classified as stressed by the actual number of stressed subjects ($n = 36$). We calculated the specificity by dividing the number of subjects correctly classified as non-stressed by the actual number of non-stressed subjects ($n = 22$). In addition, predictive accuracy, which is the number of correctly classified subjects divided by the number of total subjects ($n = 58$), was calculated for each classifier [25].

Availability of data and material The datasets used and/or analyses during the current study are available from the corresponding author on a reasonable request.

3 Results

Calibration is an important step to ensure the quality of the measured parameters in medical devices. Table 1 shows the results of the calibration experiments performed to evaluate the accuracy and stability for the measured physiological signals. The stabilities of measuring PD, PPG, GSR, and temperature were 0.03 mm, 3 mV, 0.004 MΩ, and 0.021 °C, respectively, which indicated stable baseline reading and low noise level. The accuracy of measuring PD, GSR, and temperature was 0.1 mm, 0.05 MΩ, and 0.1 °C, respectively.

Descriptive summaries of age, weight, height, BMI, and STAI scores for the 58 subjects are given in Table 2. The test-retest reliability coefficient for the STAI total score was 0.86 ($p < .001$), indicating excellent stability over time. The average scores were (38.43 ± 9.02 SD), 39.78 ± 8.73, and 78.21 ± 16.79 for the STAI-Y1, STAI-Y2, and total STAI score, respectively. The average score of STAI for females (77.41 ± 14.97) was slightly smaller than that for males (79.19 ± 19.05), but not statistically significant ($p = 0.691$). Subjects with STAI scores above the average ($\overline{STAI} = 78.21$) were considered as stressed subjects, and those with STAI scores below the average were considered as non-stressed. Twenty-two subjects (10 males and 12 females) were classified as stressed, and 36 subjects (16 males and 20 females) were classified as non-stressed. There was no significant difference in age between stressed subjects

Table 1 System performance

Type	Parameter	Value
Sampling the physiological signals	Sampling frequency	60 samples per second
Imaging the left eye	Imaging rate	15 images per second
Pupil diameter (PD)	Stability over 100 images using a phantom with 5 mm pupil diameter)	0.03 mm
	Accuracy	0.1 mm
Galvanic skin response (GSR)	Stability over 15 min using a 1-MΩ resistance	0.004 MΩ
	Accuracy	0.05 MΩ
	Linear operating range	0–2.1 MΩ
PPG	Stability over 15 min when covered by the densest slap of silicon	3 mV
Thermistor (for breathing monitoring)	Stability over 15 min at 23 °C	0.021 °C
	Accuracy	0.1 °C
	Linear operating range	20–40 °C

(21.50 ± 1.26 years) and non-stressed subject (21.52 ± 1.58 years).

Figure 5 shows an examples of normalized HR ($\acute{v}_{HR}[n]$) and GSR ($\acute{v}_{GSR}[n]$) during math questions for a 23-year-old female subject, who was classified as non-stressed subject (STAI score equal 67). Inspection of the figure reveals an increase in HR and a decrease in GSR in range of 5 to 10% during the cognitive load intervals (30–90 s). Normalization to baseline was done to reduce intersubject variabilities of physiological parameters and hence allowed comparisons of CL-induced changes in physiological parameters between subjects.

Figure 6 shows a bar graph of the average value of $\acute{v}[n]$ during cognitive-load intervals (Δ) across all subjects ($n = 58$). The average and standard deviation were 5.25 ± 4.83, 11.58 ± 19.71, 1.63 ± 8.22, − 11.11 ± 9.09, and 8.88 ± 10.75 for Δ_{HR}, Δ_{PWA}, Δ_{BR}, Δ_{GSR}, and Δ_{PD}, respectively. Note that HR, PWA, GSR, and PD were significantly changed ($p < 0.01$) during cognitive load intervals across subjects (different from

zero), while BR was not significantly changed. There were no significant correlations between changes in physiological parameters. The correlation between the total STAI scores and changes in the physiological parameters are shown in Table 3. The only significant correlation was found between Δ_{PWA} and STAI scores ($p < 0.05$), with a negative correlation coefficient (− 0.300).

Figure 7 shows the average value of $\acute{v}_x[n]$ during cognitive-load intervals (Δ) for stressed subjects ($n = 22$, red bars) versus non-stressed subjects ($n = 36$, blue bars). The average values and STD of Δ_{HR}, Δ_{PWA}, Δ_{BR}, Δ_{GSR}, and Δ_{PD} for non-stressed subjects were 5.92 ± 4.31, 16.87 ± 21.39, 0.92

Table 2 Demographic, STAI scores, and physiological parameters information of the subjects

Parameter	Min	Max	Average	Std deviation
Age (year)	19	25	21.52	1.45
Weight (kg)	41	95	64.66	12.83
Height (cm)	146	187	167.19	9.87
BMI (kg/m^2)	14.53	30.70	23.03	3.45
HR rest (bpm)	56.4	109.3	83.2	13.7
PWA rest (V)	0.23	1.64	0.82	0.37
BR rest (breath per min)	7.5	24.8	17.6	2.9
PD rest (mm)	3.25	8.41	6.32	2.67
GSR rest (MΩ)	0.26	1.95	1.04	0.51
STAI-Y1	21	59	38.43	9.02
STAI-Y2	22	61	39.78	8.73
STAI (total score)	46	120	78.21	16.79
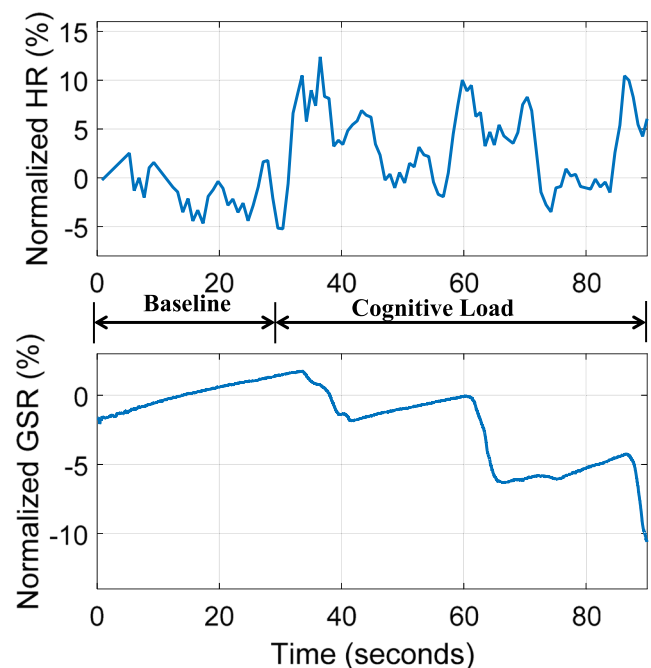
Fig. 5 An example of normalized heart rate (HR) and galvanic skin response (GSR). This data was from a 23-year-old female during baseline interval (0–30 s) and cognitive load interval (30–90 s)
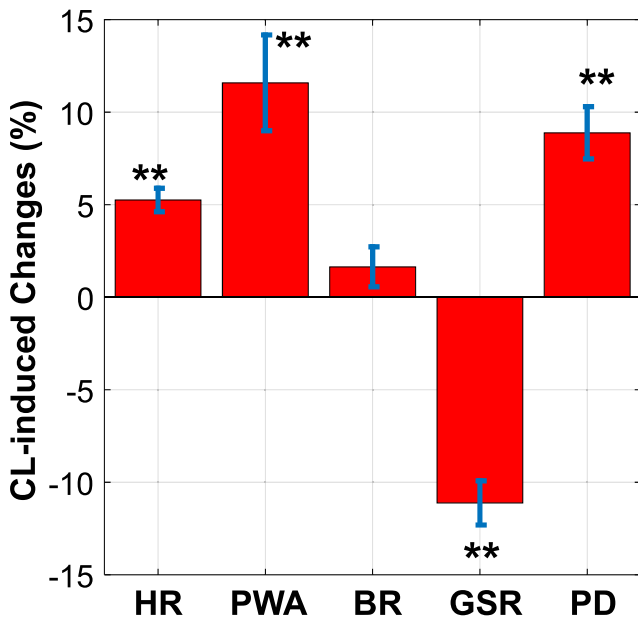
Fig. 6 The average values of changes (%) in physiological parameters for 58 subjects between cognitive load intervals and baseline intervals. Error bars: standard error. ** $p < 0.01$
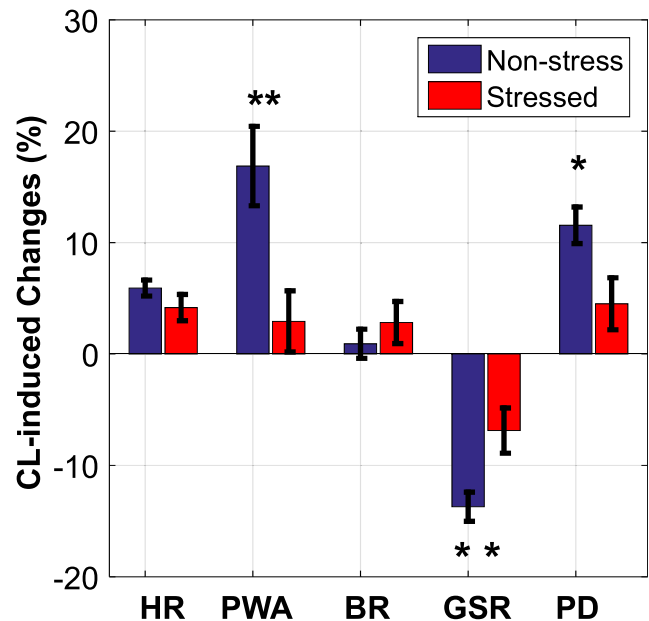

Fig. 7 The average value of CL-induced changes (percentage) in physiological parameters for stressed subjects (22 cases) versus non-stressed subjects (36 cases). Error bars: standard error. * $p < 0.05$; ** $p < 0.01$

± 7.84, -13.71 ± 7.86, and 11.55 ± 9.85, respectively, and those for stressed subjects were 4.17 ± 5.52, 2.93 ± 12.89, 2.82 ± 8.87, -6.87 ± 9.54, and 4.51 ± 10.94. Interestingly, the average values of Δ_{PWA} and Δ_{PD} for non-stressed subjects were significantly greater than that for stressed subjects ($p < 0.01$ for Δ_{PWA}, $p < 0.05$ for Δ_{PD}). Furthermore, the average values of Δ_{GSR} for non-stressed subjects were significantly less than that for stressed subjects ($p < 0.01$). In contrast, the average values of Δ_{HR} and Δ_{BR} were not significantly different between stressed and non-stressed subjects ($p = 0.183$ for Δ_{HR}, $p = 0.396$ for Δ_{BR}). These results suggest that PWA, GSR, and PD are sensitive to the presence of chronic stress and can be used to detect stressed subjects.

Figure 8 shows the area under ROC (AUC) $\pm$ standard error of the five Δs for classifying stressed subjects from non-stressed subjects. Acceptable AUC values 0.742 ± 0.078, 0.724 ± 0.071, and 0.740 ± 0.071 were obtained for Δ_{PWA}, Δ_{GSR}, and Δ_{PD}, respectively. While AUC values 0.572 ± 0.079 and 0.451 ± 0.078 for Δ_{HR} and Δ_{BR} were not significantly greater than 0.5. The ROC plots are provided in the Supplementary Material.

In this study, we applied five classification algorithms to build models that discriminated stressed subjects from non-
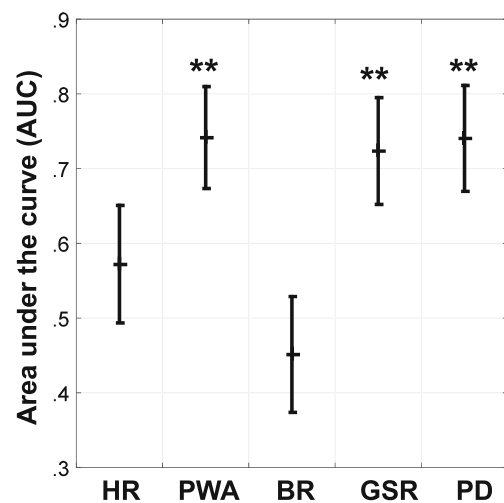
stressed subjects inputting all extracted features: Δ_{HR}, Δ_{PWA}, Δ_{BR}, Δ_{GSR}, and Δ_{PD}. The applied classifiers were logistic regression, Naïve Bayes, Multilayer Perceptron, Random forest, and K-star (K*). The goal of using these classification algorithms was to develop and train a system that accept all extracted features as input and then indicate whether an unlabeled subject has chronic stress (stressed) or not (non-stressed). The confusion matrix and predictive accuracy of the classification algorithms are shown in Table 4. The italicized values represent the number of correctly classified

Table 3 Correlation coefficients between changes in the physiological parameters of participants and their STAI total scores

Physiological parameter	Δ_{HR}	Δ_{PWA}	Δ_{BR}	Δ_{GSR}	Δ_{PD}
Correlation coefficient (ρ)	-0.049	-0.300	-0.026	0.189	-0.130
p value	0.718	0.022	0.845	0.155	0.32

Fig. 8 The area under ROC of the five physiological parameters for their ability to differentiate between stressed subjects (22 cases) from non-stressed subjects (36 cases). Results are shown as average $\pm$ standard error (SE, nonparametric bootstrap-based estimator in SPSS). ** $p < 0.01$

Table 4 Confusion matrix, sensitivity, specificity, and predictive accuracy for chronic stress detection using five classification algorithms (italicized values in the confusion matrices are the number of correctly classified states)

			Classification		Sens. (%)	Spec. (%)	Predictive accuracy (%)
			Stressed	Non-stressed			
Actual	Logistic regression	Stressed	*32*	4	88.8	77.3	84.5
		Non-stressed	5	*17*			
	Multilayer Perceptron	Stressed	*33*	3	91.7	86.4	89.7
		Non-stressed	3	*19*			
	Naïve Bayes	Stressed	*33*	3	91.7	77.3	86.2
		Non-stressed	5	*17*			
	Random forest	Stressed	*32*	4	88.9	77.3	84.5
		Non-stressed	5	*17*			
	K*	Stressed	*31*	5	86.1	86.4	86.2
		Non-stressed	3	*19*			

subjects. The confusion matrices were produced by applying classification algorithms on subjects in the validating subsamples; hence, they give an estimate of the performance of each algorithm for unseen future data. The numbers of correctly classified subjects were 49, 52, 50, 49, and 50 using the logistic regression, multilayer perceptron, Naïve Bayes, random forest, and K*, respectively. The maximum predictive accuracy of 89.7% was produced by using multilayer perceptron, while logistic regression and random forest algorithms produced the minimum predictive accuracy (84.5%). A maximum sensitivity of 91.7% was achieved by using either multilayer perceptron or Naïve Bayes, and a maximum specificity of 86.3% was achieved by using either multilayer perceptron or the K* algorithm.

4 Discussion

In the present study, a system was designed to measure CL-induced changes in physiological parameters, to identify people with chronic stress, assuming that these changes would reveal SNS reactions that were blunted by chronic stress. An easy-to-use device was designed to simultaneously measure five physiological parameters (HR, PWA, BR, GSR, and PD) innervated by the SNS, during baseline and cognitive load (number-multiplication tasks). These parameters were recorded from 58 healthy subjects, who were identified as either stressed or non-stressed, based on their STAI scores.

The merit of measuring changes in physiological parameters due to cognitive load (answering multiplying questions with different difficulty levels) is to reduce intersubject variations in physiological parameters and to reveal the autonomic nervous system (ANS) response to CL. For example, heart rates for non-stressed people ranged from 60 to 95 bpm, and heart rates for stressed people also ranged from 60 to 95 bpm.

The same hold true for other physiological parameters such as PD and GSR. Hence, absolute measuring of these parameters will less likely be helpful for predicting the presence of chronic stress, because of normal (natural) intersubject variations in these parameters. We used cognitive load (CL) to induce changes in physiological parameters that are under the control of ANS, in order to obtain relative changes to baseline. The relative changes can then be normalized to baseline values in order to reduce interpatient variations in these parameters. The magnitude of normalized changes, which depends on the ANS responses to CL, has the potential to reveal the level of chronic stress for two reasons: (1) intersubject variations are minimized or eliminated, and (2) chronic stress has been reported to impair ANS responses to CL (reduced response) [43].

The system design was optimized to increase its sensitivity and specificity to CL-induced changes in PD. The optimizations were based on recommendations from previous studies. Measuring PD while both eyes are covered by the headset (complete darkness) was reported to eliminate spontaneous variations in PD due to accommodation responses [4, 33] and to enhance CL-induced changes in PD [33]. Subjects were asked the number-multiplication questions verbally, since CL-induced changes in PD were reported to be larger when questions were asked verbally than when they were presented visually [21].

Figure 6 shows statistically significant increases in HR, PWA, and PD, and a statistically significant decrease in GSR in all subjects during CL intervals, compared to their values during baseline intervals. An increase in HR during CL was reported in several previous studies [14, 49, 53], as were an increase in PD [53, 54] and a decrease in GSR [49, 53, 54]. Minakuchi et al. [26] reported a decrease in PWA (sometimes referred to as blood volume pulse) determined from finger PPG during CL, indicating a vasoconstriction and reduction of blood flow to peripherals. PPG has also been

documented to be sensitive to increased blood flow in the frontal vessel of the forehead during CL [46].The increased PWA seen in the current study, determined by using a forehead PPG, indicated increased blood flow to the forehead. In addition, the large CL-induced changes in GSR and PD (> 8% each) indicated their sensitivity to CL, consistent with the findings of other researchers [26, 49, 54]. However, a large CL-induced increase in PWA (> 10%) determined from a forehead PPG has not been reported before. This finding increases the number of physiological parameters known to be sensitive to CL.

A controversial result was the CL-induced increase in BR, which was not statistically significant ($p > 0.05$) in this study. Minakuchi et al. [26] showed a significant BR increase during a Stroop color-word conflict test, while Wijsman et al. [49] showed an insignificant BR increase during calculations and logical puzzle tasks. The number-multiplication protocol used in the present study was similar to the calculation task used by Wijsman et al. and gave a comparable (insignificant) increase in BR during CL. Another potential reason of why BR was not a good feature for stress detection is that our techniques for measuring BR depend on the detected increase in temperature due to exhaled air from nostril, which is accurate for detecting BR if the subject during measurement keep breathing from his/her nose. Hence, breathing from mouth was not detected. Furthermore, saying answers of questions aloud during measurement might alter breathing rhythm. Indeed, the effects of CL on BR were far from being useful for stress detection [13, 47].

The results of this study provide a major step forward in monitoring chronic stress, since they showed not only that HR, PWA, GSR, and PD are sensitive to CL but also that PWA, GSR, and PD are sensitive to the presence of chronic stress that is diagnosed by STAI. Figure 7 shows significantly greater CL-induced changes in PWA, GSR, and PD in non-stressed subjects than in stressed subjects. In addition, the AUCs for PWA, GSR, and PD were significantly larger than 0.5 (Fig. 8), indicating the acceptability of each physiological feature for use in diagnosing chronic stress. In addition, predictive accuracies of using different classification algorithms (see Table 4) to combine the diagnostic power of all features yielded high levels (up to 89.7%) of differentiation between stressed and non-stressed subjects. Of the five classification algorithms, multilayer perceptron produced the highest sensitivity, specificity, and predictive accuracy. Multilayer perceptron has been shown to outperform other classifiers for many medical applications, since it uses a learning algorithm to find the best relationship between input and output variables [7, 12].

Indeed, the reported information of the classifiers in Table 4 were calculated using 10-fold cross-validation, where each data point has the chance of being in the training and validation set [6]. Particularly, these accuracies are more realistic than using all data for training, and they give an estimation of the performance of our system for unseen future data. No other studies have compared CL-induced changes in physiological parameters in people classified as stressed or non-stressed by their STAI scores.

Since variations in PWA, PD, and GSR are known to reveal SNS reaction to cognitive loads [26, 33, 54], the decreased variations in these parameters in stressed subjects suggested that the SNS mechanisms controlling these parameters were blunted. An explanation of this result is that CL in non-stressed subjects induces SNS reactions from rest into active states, while CL in stressed subjects induces SNS reactions from active to slightly more active states. This is seen in stressed subjects as reduced variations in the physiological parameters.

Teixeira et al. [43] studied CL-induced variation in saliva alpha-amylase activity, which is well accepted as a surrogate biomarker of ANS activity. In their study, high concentrations of cortisol and high scores of perceived stress were used to determine which subjects were stressed, and a Stroop test was used as a CL. They investigated CL-induced variations in saliva alpha-amylase activity between stressed and non-stressed subjects. They observed an increase in saliva alpha-amylase activity after CL in non-stressed males. This increase was not observed in stressed males or in females regardless of their stress state. Their results indicated that chronic stress leads to a hypo-activity of the ANS during CL and described the blunted response of the ANS in stressed males. From studying the ANS response to CL by measuring physiological parameters, which is easier to perform, our results indicated that people with chronic stress have a blunted reactivity of the ANS, regardless of their gender.

A limitation of the research design was that both eyes were covered by the headset during the measurement, which was inconvenient for some subjects (13 subjects). The weight of the headset was 0.330 kg (0.73 lb). After mounting the headset on the head using the lateral and medial bands, it was considered lightweight from subjects' point of view. The short time period of the measurement (< 2 min) relieved its inconvenience. We were aware that the headset might induce stress to volunteers during measurement. However, any possible induced stresses should be the same for all subjects and that minimize its effect on measuring CL-induced changes in physiological parameters. In contrast, covering both eyes by the headset might help creating constant experimental conditions for all subjects by preventing external light (room light) from inducing large changes in pupil diameter. For improvement, the design of the headset can be made lighter by placing the Raspberry Pi board beside the DAQ (not mounted on the headset), and that will make it more convenient for subjects during measurement.

Another limitation was the relatively small sample size, which included only young adults (19–25 years old). To

generalize the results, a larger sample with a wider range of ages (19–65 years old) should be studied. We are aware that the age range of participants in our study was narrow and included only 58 undergraduate students, but we still can compare the stressed with non-stressed among those participants because many studies reported that the depressions and anxiety affect more than 15% of undergraduate students [8, 51]. In one study, 3100 students were studied and found that nearly half of them were experiencing stress-related problems during the previous 12 months [16]. In addition, according to the National Survey of Counseling Center Directors in 2011, 91% of directors agreed that the number of students with severe psychological problems continues to rise [29]. On the other hand, the narrow age range of participants can be considered as an advantage in our study, since the measured CL-induced changes in physiological parameters are affected solely by the level of chronic stress rather than the age difference between subjects. In other words, the effect of age on the measured physiological parameters was eliminated by studying participants of narrow age range. From the previous reasons, we believe that our human study on the 58-student sample included good portion of stressed and non-stressed participants, which allowed us to reasonably compare stressed with the non-stressed subjects. For future directions, it will be important to evaluate the effect of aging on CL-induced changes in physiological parameters. In addition, future studies should investigate whether CL-induced changes in physiological parameters are sensitive to menstrual cycles.

We believe that there are more parameters that can be extracted from the measured signals using variability [13], frequency domain [13, 36], and principal component analysis [49]. Additionally, investigating more physiological signals, such as ECG, EEG, and EMG, may lead to the discovery of more parameters that are sensitive to the presence of chronic stress. However, adding these signals may make the measurements more difficult. Identifying more physiological parameters that are sensitive to chronic stress has the potential to increase our ability to early diagnose stress to help preventing long-term secondary diseases.

One of the advantages of the system described here is that performing the measurements is quite easy; the operator positions the headset and GSR sensors and makes sure they are stable, using the GUI, and then presses a "START" button. Generation of signals, image processing, and extraction of physiological parameters proceed automatically, and the results of the test can be available within seconds after the measurement. Another advantage is that the length of measurement is only 90 s, which makes it easy to conduct and comfortable for most subjects. With small modifications, the proposed device could be made low-cost, portable, and plug-and-play to be connected to any personal computer after installing a stand-alone software. This

would allow measurements to be taken anywhere, even at home, for better stress monitoring.

5 Conclusions

In the future, we may want to check our level of chronic stress regularly to prevent serious disorders resulted from late discovery of chronic stress. A customized system was presented in this study to test the applicability of physiological signals sensing, which can be easily done non-invasively, for determining a subject's chronic stress presence. Five physiological features were measured, namely galvanic skin response, pupil diameter, photoplethysmography, and breathing rate. A major challenge we encountered was large interpatient variations in these signals, which limit the ability for accurate detection of chronic stress. This limitation was overcome in this study by normalizing changes in physiological signals due to cognitive load to their values during rest.

Interestingly, normalized changes in three features (PWA, GSR, and PD) extracted from the measured physiological signals were more pronounced in non-stressed subjects than they were in subjects determined by STAI to be stressed. The values of area under the curve (AUC) for PWA, GSR, and PD were significantly higher than 0.5 illustrating the usefulness of each feature to identify chronic tress. Predictive accuracies of using five classification algorithms by inputting all features were sufficient even after 10-fold cross-validation (up to 89.7%), indicating the ability of the system to objectively diagnose chronic stress for unseen participants. In conclusion, the reduced CL-induced changes in PWA, GSR, and PD in stressed people support the hypothesis that chronic stress impairs the ANS reaction to CL. In addition, measuring the normalized changes in PWA, GSR, and PD can be easily used to detect the presence of chronic stress before it causes life-threatening disorders.

Compliance with ethical standards

Ethics approval and consent to participate Human studies were conducted in Jordan University of Science and Technology (JUST) and were approved by the institutional review board. The measurements were performed on healthy volunteers non-invasively and without any risk to the participants. Informed consent was obtained from all individual participants included in the study.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Alberdi A, Aztiria A, Basarab A (2016) Towards an automatic early stress recognition system for office environments based on multi-modal measurements: a review. J Biomed Inform 59(Feb. 2016): 49–75. https://doi.org/10.1016/j.jbi.2015.11.007

2. Barnes LLB, Harp D, Jung WS (2002) Reliability generalization of scores on the Spielberger's State-Trait Anxiety Inventory. Educ Psychol Meas 62(4):603–618. https://doi.org/10.1177/0013164402062004005

3. Bauer AM, Quas JA, Boyce WT (2002) Associations between physiological reactivity and children's behavior: advantages of a multisystem approach. J Dev Behav Pediatr 23(2):102–113

4. Beaty J, Lucero-Wagoner B (2000) The pupillary system. Handbook of Psychophysiology. Cambridge University Press, Cambridge, pp 142–162

5. Bell AC, D'Zurilla TJ (2009) The influence of social problem-solving ability on the relationship between daily stress and adjustment. Cogn Ther Res 33(5):439–448. https://doi.org/10.1007/s10608-009-9256-8

6. Braga-Neto U, Hashimoto R, Dougherty ER, Nguyen DV, Carroll RJ (2004) Is cross-validation better than resubstitution for ranking genes? Bioinformatics 20(2):253–258

7. Caravaca J, Soria-Olivas E, Bataller M, Serrano AJ, Such-Miquel L, Vila-Francés J, Guerrero JF (2014) Application of machine learning techniques to analyse the effects of physical exercise in ventricular fibrillation. Comput Biol Med 45:1–7. https://doi.org/10.1016/j.compbiomed.2013.11.008

8. Eisenberg D, Gollust SE, Golberstein E, Hefner JL (2007) Prevalence and correlates of depression, anxiety, and suicidality among university students. Am J Orthopsychiatry 77(4):534–542. https://doi.org/10.1037/0002-9432.77.4.534

9. Frank E, Hall MA, Witten IH (2016) The WEKA Workbench. Online appendix for data mining: practical machine learning tools and techniques, 4th edn. Morgan Kaufmann, San Francisco

10. Grös DF, Antony MM, Simms LJ, McCabe RE (2007) Psychometric properties of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): comparison to the State-Trait Anxiety Inventory (STAI). Psychol Assess 19(4):369–381. https://doi.org/10.1037/1040-3590.19.4.369

11. Hayashi T, Okamoto E, Nishimura H, Mizuno-Matsumoto Y, Ishii R, Ukai S (2013) Beta activities in EEG associated with emotional stress. International Journal of Intelligent Computing in Medical Sciences & Image Processing 3:57–68. https://doi.org/10.1080/1931308X.2009.10644171

12. Haykin SS and Haykin SS (2009) Neural networks and learning machines. Prentice Hall

13. Healey JA, Picard RW (2005) Detecting stress during real-world driving tasks using physiological sensors. IEEE Trans Intell Transp Syst 6(2):156–166. https://doi.org/10.1109/TITS.2005.848368

14. Hjortskov N, Garde AH, Ørbæk P, Hansen ÅM (2004) Evaluation of salivary cortisol as a biomarker of self-reported mental stress in field studies. Stress Health 20(2):91–98. https://doi.org/10.1002/smi.1000

15. Hreidarsson AB (1982) Pupil size in insulin-dependent diabetes. Relationship to duration, metabolic control, and long-term manifestations. Diabetes 31(5 Pt 1):442–448

16. Hyun J, Quinn B, Madon T, Lustig S (2006) Graduate student mental health: needs assessment and utilization of counseling services. J Coll Stud Dev 47(2006):247–266

17. ISMA (2013) Stress questionnaire. International Stress Management Association, UK

18. Kaga M, Nakamoto Y, Nakamura K, Ikeda K, Yoshii M, Kawana S (2014) Stress sensitivity in patients with atopic dermatitis in relation to the translocator protein 18 kDa (TSPO). Journal of Nippon Medical School 81:148–156. https://doi.org/10.1272/jnms.81.148

19. Kaklauskas A, Zavadskas E, Seniut M, Dzemyda G, Stankevic V, Simkevicius C, Stankevič T, Paliskiene R, Matuliauskaite A, Kildiene S, Bartkiene L, Ivanikovas S, Gribniak V (2011) Web-based biometric computer mouse advisory system to analyze a user's emotions and work productivity. Eng Appl Artif Intell 24:928–945. https://doi.org/10.1016/j.engappai.2011.04.006

20. von Känel R, Mills PJ, Fainman C, Dimsdale JE (2001) Effects of psychological stress and psychiatric disorders on blood coagulation and fibrinolysis: a biobehavioral pathway to coronary artery disease? Psychosom Med 63(4):531–544

21. Klingner J, Tversky B, Hanrahan P (2011) Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. Psychophysiology 48(3):323–332. https://doi.org/10.1111/j.1469-8986.2010.01069.x

22. Lewinsohn PM, Rohde P, Seeley JR, Fischer SA (1993) Age-cohort changes in the lifetime occurrence of depression and other mental disorders. J Abnorm Psychol 102(1):110–120

23. Lupien S and Seguin F (2013) How to measure stress in humans. Tech. rep. Centre for Studies on Human Stress, Quebec

24. Maier W, Buller R, Philipp M, Heuser I (1988) The Hamilton Anxiety Scale: reliability, validity and sensitivity to change in anxiety and depressive disorders. J Affect Disord 14(1):61–68

25. Milligan GW and Hirtle SC (2003) Clustering and classification methods. In: IB Weiner (ed) Handbook of Psychology. John Wiley & Sons, Inc.

26. Minakuchi E, Ohnishi E, Ohnishi J, Sakamoto S, Hori M, Motomura M, Hoshino J, Murakami K, Kawaguchi T (2013) Evaluation of mental stress by physiological indices derived from finger plethysmography. J Physiol Anthropol 32:17. https://doi.org/10.1186/1880-6805-32-17

27. Nakamoto, Ikeda Y, Nakamura K, Yoshii M (2016) Gender difference in stress responses as examined by TSPO (Translocator Protein 18 kDa), cortisol and STAI (State-Trait Anxiety Inventory) in normal human subjects. Int J Neuropsychopharmacol 19(Suppl_1):10–11. https://doi.org/10.1093/ijnp/pyw043.030

28. Nater UM, Rohleder N, Gaab J, Berger S, Jud A, Kirschbaum C, Ehlert U (2005) Human salivary alpha-amylase reactivity in a psychosocial stress paradigm. International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology 55(3):333–342. https://doi.org/10.1016/j.ijpsycho.2004.09.009

29. National Survey of Counseling Center Directors 2011 (2012) http://d-scholarship.pitt.edu/28174/. Accessed: 2018-03-25

30. Nemeroff CB, Goldschmidt-Clermont PJ (2012) Heartache and heartbreak—the link between depression and cardiovascular disease. Nat Rev Cardiol 9(9):526–539. https://doi.org/10.1038/nrcardio.2012.91

31. Nowak W, Żarowska A, Szul-Pietrzak E, Misiuk-Hojło M (2014) System and measurement method for binocular pupillometry to study pupil size variability. Biomed Eng Online 13:69–85. https://doi.org/10.1186/1475-925X-13-69

32. Palanisamy K, Murugappan M, Yaacob S (2013) Multiple Physiological Signal-Based Human Stress Identification Using Non-Linear Classifiers. Elektronika ir Elektrotechnika 19(7):80–85. https://doi.org/10.5755/j01.eee.19.7.2232

33. Pedrotti M, Mirzaei MA, Tedesco A, Chardonnet J-R, Mérienne F, Benedetto S, Baccino T (2014) Automatic stress classification with pupil diameter analysis. Int J Hum Comput Interact 30(3):220–236. https://doi.org/10.1080/10447318.2013.848320

34. Physiology (2015) <http://www.oxforddictionaries.com/definition/english/physiology>

35. Psychological (2015) <http://www.collinsdictionary.com/dictionary/english/psychological?showCookiePolicy=true>

36. Reiner M, Gelfeld TM (2014) Estimating mental workload through event-related fluctuations of pupil area during a task in a virtual world. Int J Psychophysiol 93(2014):38–44

37. Selesnick IW, Graber HL, Pfeil DS, Barbour RL (2014) Simultaneous low-pass filtering and total variation denoising. IEEE Trans Signal Process 62(5):1109–1124. https://doi.org/10.1109/TSP.2014.2298836

38. Selye H (1982) History and present status of the stress concept. In: Goldberger L, Breznitz S (eds) Handbook of stress: theoretical and clinical aspects. Free Press, New York, NY

39. Serlachius A, Hamer M, Wardle J (2007) Stress and weight change in university students in the United Kingdom. Physiol Behav 92(4):548–553. https://doi.org/10.1016/j.physbeh.2007.04.032

40. Spielberger CD (1984) State-trait anxiety inventory: a comprehensive bibliography. Consulting Psychologists Press

41. Stark L (1968) Neurological control systems—studies in bioengineering | Lawrence Stark | Springer. In Studies in Bioengineering. New York, Plenum Press

42. Steptoe A (1991) Invited review. The links between stress and illness. J Psychosom Res 35(6):633–644

43. Teixeira RR, Díaz MM, da Santos TVS, Bernardes JTM, Peixoto LG, Bocanegra OL, Neto MB, Espindola FS (2015) Chronic stress induces a hyporeactivity of the autonomic nervous system in response to acute mental stressor and impairs cognitive performance in business executives. PLoS ONE 10(3). https://doi.org/10.1371/journal.pone.0119025

44. Tsigos C, Chrousos GP (2002) Hypothalamic-pituitary-adrenal axis, neuroendocrine factors and stress. J Psychosom Res 53(4):865–871

45. Vese LA, Osher SJ (2004) Image denoising and decomposition with total variation minimization and oscillatory functions. Journal of Mathematical Imaging and Vision 20(1–2):7–18. https://doi.org/10.1023/B:JMIV.0000011316.54027.6a

46. Wager TD, Hernandez L, Jonides J, Lindquist M (2007) Elements of functional neuroimaging. Handbook of Psychophysiology. Cambridge University Press, Cambridge, pp 19–55

47. Wei CZ (2013) Stress emotion recognition based on RSP and EMG signals. Adv Mater Res 709(2013):827–831. https://doi.org/10.4028/www.scientific.net/AMR.709.827

48. Wendelken SM, McGrath SP, Akay M and Blike GT (2004) Using a forehead reflectance pulse oximeter to detect changes in sympathetic tone. The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp 325–328

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

50. Willis BH, Riley RD (2017) Measuring the statistical validity of summary meta-analysis and meta-regression results for use in clinical practice. Stat Med 36(21):3283–3301. https://doi.org/10.1002/sim.7372

51. Wyatt T, Oswalt SB (2013) Comparing mental health issues among undergraduate and graduate students. Am J Health Educ 44(2):96–107. https://doi.org/10.1080/19325037.2013.764248

52. Yokoyama K, Araki S, Kawakami N, Tkakeshita T (1990) Production of the Japanese edition of profile of mood states (POMS): assessment of reliability and validity. Nihon Koshu Eisei Zasshi, Japanese Journal of Public Health 37(11):913–918

53. Zhai J and Barreto A (2006) Stress detection in computer users based on digital signal processing of noninvasive physiological variables. Conference proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference. 1:1355–1358. https://doi.org/10.1109/IEMBS.2006.259421

54. Zhou J, Sun J, Chen F, Wang Y, Taib R, Khawaji A, Li Z (2015) Measurable decision making with GSR and pupillary analysis for intelligent user interface. ACM Trans Comput Hum Interact 21(6):33:1–33:23. https://doi.org/10.1145/2687924
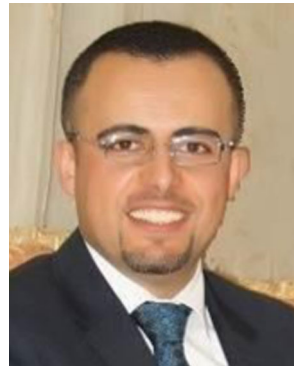
Rabah M. Al abdi received his Ph.D. degree in Biomedical Engineering from State University of New York in 2012. He is currently an assistant professor at Jordan University of Science and Technology.

Ahmad E Alhitary received his BSc degree in Biomedical Engineering from Jordan University of Science and Technology in 2016. He is currently a biomedical engineer at Al-Amal Arabia Specialized Hospital, Yemen.

Enas Abdulhay received her Ph.D. degree in Biomedical Engineering from Joseph Fourier University, France, in 2010. She is currently an assistant professor at Jordan University of Science and Technology.

Areen Al-Bashir received her Ph.D. degree in Biomedical Engineering from Wayne State University, USA in 2010. She is currently an assistant professor at Jordan University of Science and Technology.