Toward Automatic Detection of Acute Stress: Relevant Nonverbal Behaviors and Impact of Personality Traits

David Antonio Gómez Jáuregui
ESTIA
Bidart, France
Email: d.gomez@estia.fr

Carole Castanier, Bingbing Chang
CIAMS, Université Paris-Sud
Orsay, France
Email: {carole.castanier, bing-bing.chang}@u-psud.fr

Michaël Val
CIAMS, Université Paris-Sud
ESO Paris Recherche
Orsay, France
Email: michaelval@outlook.com

François Cottin, Christine Le Scanff
CIAMS, Université Paris-Sud
Orsay, France
Email: {francois.cottin, christine.le-scanff}@u-psud.fr

Jean Claude Martin
LIMSI-CNRS
Orsay, France
Email: martin@limsi.fr

***Abstract*—The aim of the present study is to identify relevant nonverbal features allowing the discrimination of different stressful behaviors, with the consideration of personality factors. In order to achieve this aim, we propose a new method for psychological stress induction involving four different stressful tasks. The proposed protocol was tested with 45 PhD students and the analysis of heart rate variability suggests that stress was indeed elicited. PhD students were selected as participants because they often experience stress. Multimodal data was collected and analyzed in order to identify nonverbal behavioral features related to the different stressful tasks. The psychological profile of participants was taken into account to understand how different stressful behaviors are correlated with personality factors. Results suggest that relevant nonverbal behaviors can discriminate between stressful tasks. In addition, relevant behaviors involving movement variability appear to be correlated with personality factors and stressful tasks.**

1. Introduction

Acute stress occurs when individuals perceive that they cannot adequately support the demands imposed on them and that pose a threat to their well-being. According to Lazarus [1], stress is a two-way process: it involves the production of stressors by the environment and the response of an individual subjected to these stressors (e.g., stress response to socially evaluative situations [2]). Studies have shown that stress represents a 50% risk factor for coronary and cardiovascular disease [3]. Correlations between stress and individual behaviors such as smoking, accident risks, absenteeism, and aggressiveness are also reported [4].

Automatic stress detection is an emerging research topic in Affective Computing research. A stress detection system could help users to better understand and manage their stress. Several systems have been developed for the detection of stressful states. These systems and their evaluation make use of the analysis of physiological signals, including blood volume pressure [5], heart rate variability (HRV) [6], skin conductance [7], and cortisol saliva samples [8]. However, most of these systems require invasive sensors that may themselves induce stress in participants. Recent advances in computer vision have led to the design of non-invasive systems capable of estimating user stress from the video analysis of facial expressions, gestures, postures, gaze and blinking, and head movements [9], [10], [11], [12]. Tasks and situations commonly used for the design and evaluation of these systems include mathematical problem solving, public speaking and global overwhelming workload. However, stress recognition remains limited in these studies, mainly because all these nonverbal behaviors do not always provide useful information for stress detection during a specific task [9]. In addition, these systems and their evaluation do not always consider individual factors such as personality and individual differences, which may impact stress-coping strategies and associated behaviors [13]. Finally, although some studies have explored the relationship between stressful nonverbal behaviors and individual factors [14], none have compared the impact of different stressful tasks on these behaviors.

In the present study, we propose a method for identifying relevant nonverbal features that allow the discrimination of psychological stress related affects and their nonverbal expression. We describe four tasks for inducing stress and the collection of multimodal behavioral data from 45 PhD students. PhD students were selected as participants because they often experience stress in several situations (thesis preparation, supervisor relationship, socioeconomic problems, etc.) during their academic cycle, and difficulty in coping with these situations can provoke them to drop out of school before graduating [15], [16]. We explain how the analysis of heart rate variability suggests that stress was indeed induced in our participants. Finally, we report the correlations between the nonverbal features and the personality of our participants. Possible applications of the findings of this study include stress management systems, job interview trainings [15].

Section 2 reviews related work regarding nonverbal features for stress detection. Section 3 explains the method that we propose. Section 4 describes the nonverbal features extracted from video recordings of participants. Section 5 presents our results.

2. Related Work

The emotional state of individuals is expressed by multiple verbal [17] and nonverbal behaviors [18]. According to Ekman [19] and Wallbott [20], emotions can be accurately decoded from facial expressions, postures, gestures, and nonverbal movements. A number of researchers seek to recognize stress related affects by automatically detecting and analyzing nonverbal behaviors. Metaxas et al. [21] propose a dynamic three-dimensional (3D) model of the face for the detection of facial behaviors that can be related to stress. They describe three major indicators of stress: eyebrow movements, asymmetric lip deformations, and teeth baring. Liao et al. [22] propose a system inferring stress from the detection of nine visual features: blinking frequency, average eye closure speed, percentage of saccadic eye movement, gaze spatial distribution, percentage of large pupil dilation, pupil ratio variation, head movement, mouth openness, and eyebrow movement. These visual features are combined with physiological cues, including heart rate, skin temperature, and electrodermal activity, and task performance, including scores, number of clicks, and finger pressure, using a Dynamic Bayesian Network. The system was tested with participants performing a mental math task and an audio task.

Some researchers propose to use the identification of primary emotions using automatic recognition of facial Action Units [19]. Das and Yamada [12] analyze the relationship between primary emotions and stress behaviors. They propose a formula for stress evaluation in terms of the percentages of blends of different emotions expressed in facial expressions. Gao et al. [11] propose to identify the stress levels of participants in a car driving task from the automatic detection of anger and disgust. Aigrain et al. [9] propose to infer stress by combining facial and body cues. Facial features are extracted from facial Action Units related to eyebrow movements, lip movements, cheek raising, nose wrinkling, chin raising, and jaw dropping. The authors analyze the following bodily behaviors and features: posture shift, head-touching with one or two hands, quantity of motion of the whole body, quantity of motion of each hand, and the quantity of motion of the head. Then, the participants' self-reported stress levels were collected and used as ground truth. The authors obtain a 77% stress detection rate by combining all extracted features using a support vector machine classifier.

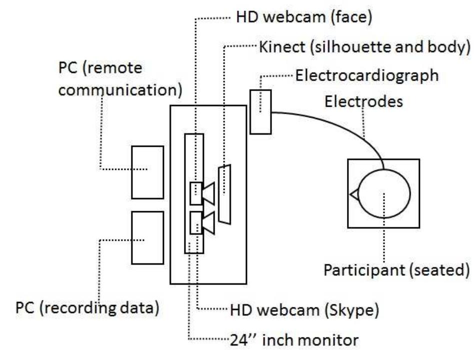

Figure 1: Top view of the set-up

Despite these advances, the automatic detection of stress from nonverbal behaviors remains limited, notably by the absence of protocols that correctly induce psychological stress. In addition, not all detected nonverbal behaviors provide relevant or useful information for stress identification. Finally, existing studies do not consider how individual differences and personality traits impact the multimodal expression of stress and, consequently, its automatic detection.

3. Method

3.1. Participants

45 PhD candidates (aged 24 to 36 years ; 15 female) were selected to participate in this study. All participants were volunteers and signed an informed consent form. The experimental protocol was approved by a local Ethics Committee. Participants were blind to the tasks or the purpose of the experiment. They were only told that they were participating in a study about PhD students.

3.2. Apparatus

The following equipment was placed in a soundproof room (Fig. 1): a PowerLab 26T (LTS) electrocardiograph with three electrodes ; a 24-inch monitor ; 2 HD Webcam Logitech Pro cameras, one to monitor the participant's behavior from the experimenter room and one to capture facial expressions of the participant ; a MS Kinect to collect the silhouette and RGB video of participants' body behaviors ; 2 desktop PCs, one to establish a Skype session with the participant and one to collect multimodal data from the participant ; two loudspeakers ; a table for placing the Kinect and loudspeakers ; a chair placed 1 m from the table and 1.1 m from the computer display and Kinect sensor. The following software were used in the study: Skype [23] ; Labchart software [24] to record and synchronize electrocardiography (ECG) and video data ; Social Signal Interpretation Framework [25] to record and process Kinect data ; Client/server application to launch audio files on a remote PC ; remote control software (Windows Remote Assistance).

3.3. Procedure

The design of our experimental protocol was inspired by experimental protocols known to elicit social stress [2] (Table 1). Before the day of the experiment, participants were asked to complete an online version of the Big Five Personality inventory [26]. On the day of the experiment, after signing the consent letter, the participant was asked to enter the soundproof room and sit on the chair placed in front of the computer. Three electrodes were attached over the right clavicle, left clavicle, and lower left abdomen [27]. The experimenter started recording multimodal data (ECG, HD webcam, and Kinect). The participant was asked to complete a one-item self-stress assessment in a paper-based questionnaire and the French version of the StateTrait Anxiety Inventory (STAI) [28]. The self-stress assessment included the following question: "At this precise moment, how stressed are you?" The participant had to provide a rating on a scale of 1 to 10. Then, the experimenter left the room and the neutral task began. The participant was instructed to inhale and exhale at a frequency of 0.25 Hz for 5 min. The participant was guided by an audio tape providing instructions regarding the precise timing of inhaling and exhaling. The participant was monitored during the task; in case of improper breathing, the task was interrupted and restarted. The purpose of this neutral breathing task was to record the resting heart rate for each participant. After the 5-min breathing exercise, the participant was asked to complete the self-stress assessment and STAI again. Then, the experimenter left the room and the stressful task began. During this task, the participant was told that an interviewer would communicate with him/her via Skype using audio only. No video of the interviewer was displayed to the participant. In reality, the interviewer was a previously recorded voice of a male University professor. To make the participant believe that he/she was actually communicating with a real person during the task, each preregistered audio file was remotely played by the experimenter from another room while a Skype audio call was displayed on the screen. The stressful task comprised the following four subtasks.

1) **Instructions:** A research laboratory director (preregistered audio) instructed the participants to present their thesis work in 3 min, including a progress report that would be evaluated by an expert committee. The participant was given 1 min to prepare the presentation.
2) **Preparation:** During the 1-min preparation, the participant was not interrupted. After 1 min, the research director interrupted and instructed the participant to begin the presentation after a beep sound.
3) **Presentation:** For this task, three actions were performed to induce more stress in the participant during the presentation. First, the tick tock sound of a clock was played from loudspeakers in the background during the 3-min presentation. Second, at 60s after the start of the presentation, the research director interrupted the participant and asked him/her to rephrase the thesis subject. Finally, if the participant completed his presentation before 3 min, the research director asked him/her to continue until time ran out.
4) **Questions:** Once the presentation was complete, the research director asked the following questions (in the following order) to the participant: How long do you think you require to complete your thesis? On a scale of 1 to 10, how would you rate your performance during the task? On a scale of 1 to 10, how do you rate your stress level during the task?

TABLE 1: MAIN STEPS OF THE STRESS INDUCTION PROTOCOL

Time	Content
Before the experiment	Personality questionnaires
Participant arrival	Arrival in the soundproof room Initiation of multimodal data recording (ECG, Kinect, Webcam)
Before task	Questionnaires (self-reported stress assessment and STAI)
Neutral task	Five minutes of inhaling and exhaling
After the neutral task	Questionnaires (self-stress assessment and STAI)
Stressful task	Thirty seconds of instructions One minute of preparation Three minutes of presentation One minute of questions
After the stressful task	Questionnaires (self-stress assessment and STAI)
Three minutes after the stressful task	Questionnaires (self-stress assessment and STAI)

Subsequently, the research director thanked the participant and invited him/her to wait for the experimenter. Immediately after the stressful task was complete, the participant was asked to complete the self-stress assessment and STAI for a third time. Then, after 3 min, the experimenter asked the participant to complete the same questionnaires again for the last time. Finally, the participant was thanked for his/her participation and invited to leave the room after being debriefed by the experimenter in order to ensure he/she left the laboratory in a positive state. During the debriefing process, the participant was offered a snack and tea and informed about the purpose of the study. A document with recommendations regarding stress management during the PhD course was provided.

3.4. Collected data

For each participant, the following data were collected during the neutral and stressful tasks (Fig. 2): RGB video of the face (captured by webcam) ; Image resolution: 800 600 at 30 fps ; RGB video of the upper body (captured by Kinect) Image resolution: 640 480 at 30 fps ; ECG signals ; Extracted silhouette from the upper body (captured by Kinect). In addition, the following questionnaires were collected during the experiment: four self-stress assessment questionnaires (scale of 1 to 10) ; four STAIs ; one Big

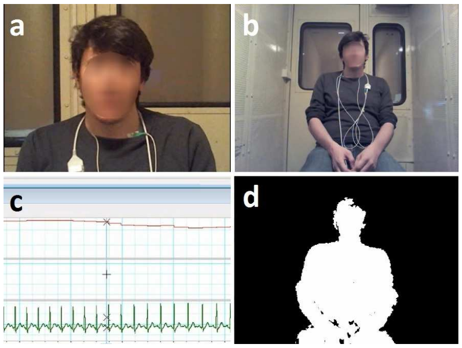

Figure 2: RGB video of face (a), RGB video of the upper body (b), ECG recording (c), and silhouette extracted from the upper body (d)

Five Personality Test. The full dataset cannot be publicly available for ethical and privacy reasons. We are considering delivering computed features and raw data that does not allow to identify individual participants.

4. Nonverbal Behaviors

Nonverbal behavioral features of the face and body were selected after a careful study of the collected videos. These features were extracted using the algorithms described below.

4.1. Facial features

4.1.1. Facial Action Units. The Facial Action Coding System (FACS) decomposes facial behaviors into AUs [19]. Each AU represents the contraction of one or several facial muscles. The activation of 19 AUs was estimated using the commercially available software FaceReader 6.0 [29]. Studies have reported a facial expression recognition accuracy of 89% with FaceReader [30], [31]. Our collected video sequences were captured under similar controlled conditions as the video sequences evaluated in [30]. The following 19 AUs were considered: Inner Brow Raiser (AU1), Outer Brow Raiser (AU2), Brow Lowerer (AU4), Upper Lid Raiser (AU5), Cheek Raiser (AU6), Lid Tightener (AU7), Nose Wrinkler (AU9), Upper Lip Raiser (AU10), Lip Corner Puller (AU12), Dimpler (AU14), Lip Corner Depressor (AU15), Chin Raiser (AU17), Lip Puckerer (AU18), Lip Stretcher (AU20), Lip Tightener (AU23), Lip Pressor (AU24), Lips Part (AU25), Jaws Drop (AU26), and Mouth Stretch (AU27). For each extracted AU, activation percentages were obtained for the neutral and stressful tasks for all participants.

4.1.2. Quantity of Motion (QoM) of the face. The quantity of motion of the face was computed (Fig. 3e). Motion history images (MHI) have been proven to be very robust in detecting motion and is widely employed by various research groups for action recognition and motion analysis [9], [10]. Other processing techniques such as optical flow or dense face tracking will be considered in future work. For each image, the face was detected using Viola-Jones face detector [32]. Then, a motion history image (MHI) [33] was computed only inside the region detected for the face. Finally, the quantity of motion was normalized with respect to the region of the face using the following equation:

$$QoM_F = \frac{M_F}{A_F} \tag{1}$$

where QoM_F is the amount of motion of the face, ranging from 0 to 1, M_F is the number of pixels in the region where motion has been detected, and A_F corresponds to the area of the face region.

4.2. Body features

4.2.1. Variability of approach and avoidance behaviors. Approach and avoidance behaviors can be estimated using the interocular distance [34]. For each image, the face and eyes are detected using Viola-Jones face detector and Haar-like features [32]. The interocular distance is obtained by computing the distance (number of pixels) between the center of the eyes. If this interocular distance increases with respect to the previous image, the system interprets this increase as an approach behavior (Fig. 3a). If the interocular distance decreases, the system interprets this decrease as an avoidance behavior (Fig. 3b). To obtain the variability of approach and avoidance behaviors of the participants, the standard deviation of the interocular distance σ_{AA} was computed using the following equation:

$$\sigma_{AA} = \sqrt{\frac{\sum_1^N \left(I_D - \tilde{I_D}\right)^2}{N}} \tag{2}$$

where I_D is the interocular distance (in pixel units) extracted from a given image and $\tilde{I_D}$ is the average value of the interocular distance from N images.

4.2.2. Variability of head orientation. FaceReader [29] tracks the orientation of the head for each image sequence. Three rotation angles are extracted using a 3D model of the face: roll, pitch, and yaw. The variability of the head orientation is defined by computing the standard deviation for each orientation angle (roll, pitch, and yaw) from a sequence of N images.

4.2.3. Contraction Index (CI). This feature indicates the degree of contraction of the body. It is represented as a value ranging from 0 to 1. A larger CI (close to 1) indicates a more closed participant posture (Fig. 3c), whereas a smaller CI (close to 0) indicates a more open participant posture (Fig. 3d). CI is computed using the following equation:

$$CI = \frac{A_S}{C_S} \tag{3}$$

where A_S is the area (in pixel units) of the extracted silhouette of the participant and C_S is the area of the bounding box surrounding the silhouette.

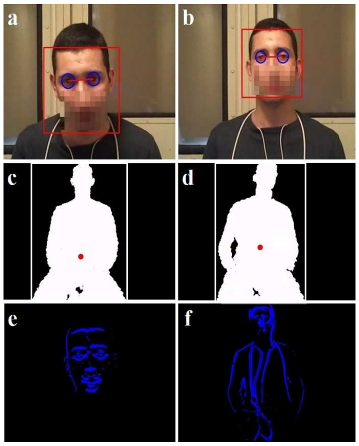

Figure 3: Extracted nonverbal features. Approach behavior (a) and avoidance behavior (b) measured by the interocular distance (red line). Closed participant posture (c) indicated by a larger CI value compared with a more open posture indicated by a smaller CI value (d). COG (red circle) vertical displacement in (d) compared with the COG (red circle) position in (c). QoM of the face (e) and the body (f)

4.2.4. Quantity of Motion (QoM) of the whole body. This is a value ranging from 0 to 1 that estimates the total amount of motion of the whole body. It is computed by extracting a motion history image (MHI) [33] from the whole body. Then, the quantity of motion QoM_B is normalized using the following equation:

$$QoM_B = \frac{M_B}{C_S} \tag{4}$$

M_B is the number of pixels in the region where motion has been detected and C_S is the area of the bounding box surrounding the silhouette of the participant's body.

4.2.5. Variability of the center of gravity displacement. The center of gravity (COG) displacement is an indicator of balance and postural control (Fig. 3c and 3d). Higher COG displacements have been reported to be associated with negative emotions and stressful situation appraisals [35]. The COG in the horizontalvertical plane was extracted directly from the participant's silhouette using the following equations:

$$X_g = \frac{\sum_i^N X_i}{N}, \quad Y_g = \frac{\sum_i^N Y_i}{N} \tag{5}$$

where X_g,Y_g is the two-dimensional COG point, X_i,Y_i is the 2D silhouette point, and N is the number of pixels in the silhouette. The variability of the COG displacement was obtained by computing the standard deviation of the COG point (X_g,Y_g) from N images.

5. Analysis and Results

Our hypotheses were as follows:

TABLE 2: PERCENTAGE OF CORRECTLY DETECTED NONVERBAL FEATURES IN THE COLLECTED VIDEO SEQUENCES

Nonverbal Feature	Accuracy rate
Facial Action Units	96.8%
QoM of the face	100%
AA (interocular distance)	97.2%
Head orientation	98.3%
Contraction Index	100%
QoM whole body	100%
COG variability	100%

QoM, Quantity of Motion; AA, approach and avoidance; HO, head orientation; COG, center of gravity displacement.

- **H1.1:** Self-stress assessment scores and sympathetic activity are significantly higher in the stressful task [6].
- **H1.2:** Sympathetic activity is significantly different between stressful subtasks.
- **H1.3:** Relevant nonverbal features allow discrimination between neutral and stressful tasks and between different stressful subtasks.
- **H1.4:** Relevant nonverbal features during the stressful subtask are correlated with personality traits.

5.1. Accuracy rate

Table 2 shows the accuracy rate of the nonverbal features extracted in our collected face and body video sequences. For all video sequences, the accuracy rate A was computed using the following equation:

$$A = \frac{N_S - N_F}{N_S} \quad X \quad 100 \tag{6}$$

where N_S is the total number of images in which the nonverbal feature was correctly extracted. N_F is the total number of images in which the extraction failed.

5.2. Elicitation of stress

5.2.1. Subjective questionnaires. An ANOVA revealed significant higher mean STAI [F(3,132) = 28.548, $p < 0.0001$] and self-stress scores [F(3,132) = 41.752, $p < 0.0001$] immediately after the stressful task than at the other points.

5.2.2. HRV analysis. Spectral analysis of HRV is widely used as a quantitative measure of stress [6]. ECG recordings from each participant were processed to extract RR interval time series. Then, smoothed pseudo WignerVille distribution (SPWVD) was used to obtain a low frequency (LF) component between 0.04 and 0.15 Hz and a high frequency (HF) component between 0.15 and 0.4 Hz. The HF component is considered to provide an estimation of the parasympathetic cardiac control, while the LF component is considered to reflect both sympathetic and parasympathetic cardiac control. Thus LF/HF ratio might reflect the sympathetic cardiac control at rest. Stress is associated with a high LF component

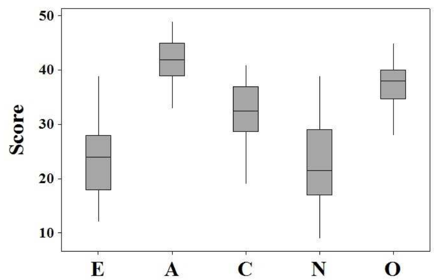

Figure 4: Boxplot of the five personality traits mean scores from all participants: extraversion (E), agreeableness (A), conscientiousness (C), neuroticism (N), openness to experience (O)

with respect to the HF component (high LF/HF ratio). The LF and HF components are normalized and expressed as a percentage of the total short-term spectral energy (LF + HF) [36]. LF/HF ratios for the neutral and stressful tasks were computed. An ANOVA showed a significant increase ($p < 0.001$) in the LF/HF ratio for the stressful task than for the neutral task. The LF component was 53.86% for the neutral task and 71.06% for the stressful task. LF/HF ratios were also obtained for each of the four subtasks. An ANOVA revealed no significant main differences ($p = 0.151$).

5.3. Analysis of Nonverbal Behavioral Features

For each feature, two one-factor within-subject ANOVA tests were conducted. In the first test, the independent variable (factor) was the type of the task, i.e., neutral or stressful. In the second test, the independent variable (factor) was the subtask, i.e., instructions, preparation, presentation, and questions. For post hoc comparisons, Bonferroni comparisons were applied. ANOVA analysis results are presented in Table 3 (AUs with no significant differences in both factors are omitted for space reasons). An analysis of effect size (Cohens d) for each nonverbal feature was also conducted in order to identify large effect sizes (d > 0.8) between two tasks (neutral vs stressful) or two stressful subtasks (e.g. instruction vs presentation). Nonverbal features with large effect sizes are presented in Table 4.

5.4. Personality Traits and Behaviors

Pearson's correlation analysis was performed to explore the relations between the nonverbal behaviors for each subtask and the big five personality traits of the participants (Fig. 4). We considered only the nonverbal behavioral features that showed significant differences ($p < 0.05$) between neutral and stressful tasks in the analyses described in the previous section. Significant correlations are shown in Table 5.

TABLE 3: ANOVA ANALYSIS RESULTS FROM EACH NONVERBAL FEATURE WITH EACH FACTOR. POST HOC COMPARISONS RESULTS ARE SHOWN IN THE FOOTNOTE REFERRED BY EACH SIGNIFICANT P-VALUE (HIGHLIGHTED IN BOLD)

Nonverbal Feature	Factor	F-value	P-value
Brow Lowerer (AU4)	task	5.93[1]	**0.018**[A]
Brow Lowerer (AU4)	subtask	0.26[2]	0.857
Cheek Raiser (AU6)	task	0.83[1]	0.366
Cheek Raiser (AU6)	subtask	6.08[2]	**0.001**[C,F]
Lip Corner Puller (AU12)	task	6.97[1]	**0.010**[B]
Lip Corner Puller (AU12)	subtask	4.49[2]	**0.005**[C,F]
Lip Puckerer (AU18)	task	5.59[1]	**0.021**[B]
Lip Puckerer (AU18)	subtask	0.43[2]	0.734
Lip Tightener (AU23)	task	6.17[1]	**0.015**[A]
Lip Tightener (AU23)	subtask	6.97[2]	**0.000**[C,D]
Lip Pressor (AU24)	task	0.90[1]	0.345
Lip Pressor (AU24)	subtask	8.09[2]	**0.000**[C,D]
Lips Part (AU25)	task	19.32[1]	**0.000**[B]
Lips Part (AU25)	subtask	12.14[2]	**0.000**[E]
Jaws Drop (AU26)	task	7.56[1]	**0.008**[B]
Jaws Drop (AU26)	subtask	5.25[2]	**0.002**[E]
QoM of the face	task	46.11[1]	**0.000**[B]
QoM of the face	subtask	24.22[2]	**0.000**[E]
AA variability	task	9.27[1]	**0.003**[B]
AA variability	subtask	4.17[2]	**0.007**[D,E,F]
HO roll variability	task	6.12[1]	**0.016**[B]
HO roll variability	subtask	6.76[2]	**0.000**[E,F]
HO pitch variability	task	0.16[1]	0.691
HO pitch variability	subtask	0.85[2]	0.469
HO yaw variability	task	7.01[1]	**0.010**[B]
HO yaw variability	subtask	10.14[2]	**0.000**[D,E,F]
Contraction Index	task	1.18[1]	0.281
Contraction Index	subtask	0.28[2]	0.842
QoM whole body	task	51.54[1]	**0.000**[B]
QoM whole body	subtask	20.33[2]	**0.000**[E]
COG (X) variability	task	4.94[1]	**0.030**[B]
COG (X) variability	subtask	1.65[2]	0.179
COG (Y) variability	task	12.97[1]	**0.001**[B]
COG (Y) variability	subtask	5.78[2]	**0.001**[D,E]

[1] F(1,44) [2] F(3,132)
[A] higher significant means for neutral task
[B] higher significant means for stressful task
[C] higher significant means for instruction subtask
[D] higher significant means for preparation subtask
[E] higher significant means for presentation subtask
[F] higher significant means for questions subtask

QoM, Quantity of Motion; AA, approach and avoidance; HO, head orientation; COG (X), variability of the horizontal center of gravity displacement ; COG (Y), variability of the vertical center of gravity displacement

6. Discussion

Subjective questionnaires and spectral analysis of HRV suggest that stress was correctly induced, supporting **H1.1**. HRV analysis revealed that the same stress level was maintained during the four stressful subtasks, these results runs contrary to **H1.2**.

The analysis of nonverbal behaviors revealed that several features are relevant for identifying stress. Most of these features can be used to discriminate between different stressful behaviors. These include Lip Corner Puller (AU12), Lip

TABLE 4: NONVERBAL FEATURES WITH LARGE EFFECT SIZES (d > 0.8) OBTAINED BETWEEN EACH TASK OR STRESSFUL SUBTASK

Nonverbal Feature	**Task/subtask (Mean ± SD)**	**Task/subtask (Mean ± SD)**	**d**
Lip Tightener (AU23)	I(23.69±29.37)	PR(3.75±7.58)	0.93
Lip Tightener (AU23)	PR(3.75±7.58)	PE(21.1±27.81)	0.85
Lip Pressor (AU24)	I(20.63±22.39)	PR(3.92±9.48)	0.97
Lip Pressor (AU24)	PE(20.29±25.8)	PR(3.92±9.48)	0.84
Lips Part (AU25)	N(10.98±20.86)	S(33.81±22.56)	1.05
Lips Part (AU25)	PR(43.03±31.7)	I(15.97±25.45)	0.94
Lips Part (AU25)	PR(43.03±31.7)	PE(15.25±21.5)	1.02
QoM of the face	N(0.009±0.007)	S(0.04±0.03)	1.67
QoM of the face	PR(0.06±0.04)	I(0.02±0.02)	1.32
QoM of the face	PE(0.02±0.02)	PR(0.06±0.04)	1.51
AA variability	I(3.77±1.78)	PR(5.22±1.8)	0.81
HO roll variability	PR(4.87±2.64)	I(2.54±2.36)	0.93
HO yaw variability	PR(4.49±2.11)	I(2.07±1.85)	1.22
HO yaw variability	I(2.07±1.85)	Q(4.27±2.7)	0.95
QoM whole body	N(0.003±0.002)	S(0.02±0.02)	1.77
QoM whole body	PR(0.03±0.02)	I(0.02±0.01)	0.84
QoM whole body	PR(0.03±0.02)	PE(0.01±0.01)	1.38
QoM whole body	PR(0.03±0.02)	Q(0.01±0.01)	1.01
COG (Y) variability	N(4.01±3.63)	S(7.55±4.14)	0.89
COG (Y) variability	PR(5.22±3.33)	Q(2.99±1.91)	0.82

N, neutral task; S, stressful task; I, instruction subtask; PR, presentation subtask; PE, preparation subtask: Q, questions subtask

QoM, Quantity of Motion; AA, approach and avoidance; HO, head orientation; COG (Y), variability of the vertical center of gravity displacement

TABLE 5: SIGNIFICANT CORRELATIONS BETWEEN NON-VERBAL BEHAVIORS AND PERSONALITY TRAITS

Nonverbal Feature	**Subtask**	**Trait**	**r-value**
QoM of the face	Preparation	E	0.333[1]
QoM of the face	Presentation	E	0.472[2]
QoM of the face	Questions	E	0.426[2]
AA variability	Presentation	C	-0.364[1]
HO roll variability	Preparation	E	0.353[1]
HO yaw variability	Presentation	C	-0.306[1]
HO roll variability	Presentation	C	-0.360[1]
HO roll variability	Questions	C	-0.466[2]
Lips Part (AU25)	Questions	A	0.327[1]
Lips Part (AU25)	Questions	N	-0.396[2]
QoM whole body	Presentation	A	-0.303[1]
QoM whole body	Questions	N	0.299[1]
QoM whole body	Questions	O	-0.303[1]
COG (Y) variability	Presentation	A	-0.323[1]
COG (Y) variability	Questions	N	0.340[1]

[1] $p < 0.05$ [2] $p < 0.01$

QoM, Quantity of Motion; AA, approach and avoidance; HO, head orientation; COG (Y), variability of the vertical center of gravity displacement ; O, openness to experience ; C, conscientiousness ; E, extraversion ; A, agreeableness ; N, neuroticism

Tightener (AU23), Lips Part (AU25), Jaws Drop (AU26), variability of approach and avoidance behaviors, QoM of the face and whole body, variability of the yaw and roll head orientation angles, and variability of vertical COG displacement. Our results suggest that relevant nonverbal features can be classified as poorly ($p < 0.05$), moderately ($p < 0.01$), and highly relevant ($p < 0.001$). In addition most of these highly relevant features proved to be highly discriminative ($d > 0.8$). However, most of these features can discriminate only between one or two stressful subtasks. Therefore, they must be combined for the automatic identification of context-dependent stress expression. These results are consistent with **H1.3**.

Correlation analyses revealed that several relevant nonverbal features were moderately correlated with different personality traits of the participants only in the preparation, presentation, and questions subtasks. Low movement variability was correlated with conscientiousness, agreeableness, and openness to experience, while high movement variability was correlated with extraversion and neuroticism. These results support **H1.4**. Knowledge of an individual's personality traits may thus be helpful for the prediction of nonverbal behaviors expressed by users during stressful tasks.

Since this study population involves only PhD students, the associated nonverbal features that we found cannot be generalized to other evaluation situational contexts (e.g. job interview for other activities). However this study suggests that, in an evaluative situation, several nonverbal features and individuals personality traits might be considered for the identification of specific stressful behaviors. Future work include designing a system for automatic stress detection. This will require data collection from a larger number of participants and integration of rules to consider personality. Speech should also be considered [37] as well as other data (thermal infrared imaging, skin conductance). Finally, the association between relevant nonverbal features and other modalities (e.g. physiological signals, perceived stress) will be also studied and modeled.

References

[1] R. S. Lazarus and S. Folkman, *Stress, Appraisal, and Coping*. New York: Springer, 1984.

[2] C. Kirschbaum, K. M. Pirke, and D. H. Hellhammer, "The 'trier social stress test'–a tool for investigating psychobiological stress responses in a laboratory setting," *Neuropsychobiology*, vol. 28, no. 12, pp. 76–81, 1993.

[3] M. Kivimaki, M. Virtanen, M. Elovainio, A. Kouvonen, A. Väänänen, and J. Vahtera, "Work stress in the etiology of coronary heart disease – a meta-analysis," *Scand. J. Work. Environ. Health*, vol. 32, no. 6, pp. 431–442, 2006.

[4] R. M. Steinmann, *Santé psychique Stress: Bases scientifiques pour une stratgie nationale en matire de prévention du stress et de promotion de la sant psychique en Suisse*. Berne et Lausanne: Promotion Santé Suisse, 2006.

[5] A. Barreto and J. Zhai, "Stress detection in computer users based on digital signal processing of noninvasive physiological variables," in *28th Annual Int. Conf. of the IEEE EMBS '06*, 2006, pp. 1355–1358.

[6] R. Bailón, L. Mainardi, M. Orini, L. Sornmo, and P. Laguna, "Analysis of heart rate variability during exercise stress testing using respiratory information," *Biomed. Signal Process. Control*, vol. 5, pp. 299–310, 2010.

[7] C. Holmgard, G. N. Yannakakis, H. P. Martinez, and K.-I. Karstoft, "To rank or to classify? annotating stress for reliable ptsd profiling," in *ACII 2015*, 2015, pp. 719–725.

[8] S. S. Dickerson and M. E. Kemeny, "Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research," *Psychol. Bull.*, vol. 130, no. 3, pp. 355–391, 2004.

[9] J. Aigrain, S. Dubuisson, M. Detyniecki, and M. Chetouani, "Person-specific behavioural features for automatic stress detection," in *11th IEEE Conf. on Automatic Face and Gesture Recognition*, 2015, pp. 1–6.

[10] J. Aigrain, M. Spodenkiewicz, S. Dubuisson, M. Detyniecki, D. Cohen, and M. Chetouani, "Multimodal stress detection from multiple assessments," *IEEE Trans. Affect. Comput.*, vol. PP, no. 99, 2016.

[11] H. Gao, A. Yuce, and J. P. Thiran, "Detecting emotional stress from facial expressions for driving safety," in *IEEE Conf. on Image Proc. (ICIP)*, 2014, pp. 5961–5965.

[12] S. Das and K. Yamada, "Evaluating instantaneous psychological stress from emotional composition of a facial expression," *Adv. Comput. Intell. Inform.*, vol. 17, p. 4, 2013.

[13] P. G. Leandro and M. D. Castillo, "Coping with stress and its relationship with personality dimensions, anxiety, and depression," *Soc. Behav. Sci.*, vol. 5, pp. 1562–1573, 2010.

[14] T. Giraud, M. Soury, J. Hua, A. Delaborde, M. Tahon, D. A. G. Jáuregui, M. V. Eyharabide, E. Filaire, C. L. Scanff, L. Devillers, B. Isableu, and J.-C. Martin, "Multimodal expressions of stress during a public speaking task," in *Affective Computing and Intelligent Interaction*, Geneva, 2013, pp. 417–422.

[15] L. Bazrafkan, N. Shokrpour, A. Yousefi, and N. Yamani, "Management of stress and anxiety among phd students during thesis writing: a qualitative study," *Health Care Manag. Frederick*, vol. 35, no. 3, pp. 231–240, 2016.

[16] N. Jaques, S. Taylor, A. Azaria, A. Ghandeharioun, A. Sano, and R. Picard, "Predicting students' happiness from physiology, phone, mobility, and behavioral data," in *ACII*, 2015, pp. 222–228.

[17] S. R. Fussell, *The Verbal Communication of Emotions: Interdisciplinary Perspectives*. Lawrence Erlbaum Associates, 2002.

[18] M. L. Knapp, J. A. Hall, and T. G. Horgan, *Nonverbal Communication in Human Interaction*. 8th edition. Wadsworth Publishing Co Inc., 2012.

[19] P. Ekman, W. V. Friesen, and J. C. Hager, *Facial action coding system: The manual. Facial action coding system, user & investigator guides*. A Human Face, 2002.

[20] H. G. Wallbott, "Bodily expression of emotion," *Eur. J. Soc. Psychol.*, vol. 28, no. 6, pp. 879–896, 1998.

[21] D. Metaxas, S. Venkataraman, and C. Vogler, "Image-based stress recognition using a model-based dynamic face tracking system," in *ICCS*, 2004, pp. 813–821.

[22] W. Liao, W. Zhang, Z. Zhu, and Q. Ji, "A real-time human stress monitoring system using dynamic bayesian network," in *CVPR*, 2005, p. 70.

[23] (2017) Skype for windows. [Online]. Available: https://www.skype.com/

[24] (2017) Labchart - adinstruments. [Online]. Available: https://www.adinstruments.com/products/labchart

[25] J. Wagner, F. Lingenfelser, T. Baur, I. Damian, F. Kistler, and E. André, "The social signal interpretation (ssi) framework: multimodal signal processing and recognition in real-time," in *ACM Multimedia*, 2013, pp. 831–834.

[26] O. Plaisant, "Validation par analyse factorielle du big five inventory franais (bfi-fr) analyse convergente avec le neo-pi-r," *Ann. Med. Psychol.*, vol. 168, pp. 97–106, 2010.

[27] G. M. Khan, "A new electrode placement method for obtaining 12-lead ecgs," *Open Heart*, vol. 2, 2015.

[28] M. Bruchon-Schweitzer and I. Paulhan, *Le manuel du STAI-Y de C.D. Spielberger*. Paris, France: ECPA, 1993.

[29] (2014) Facereader: Tool for automatic analysis of facial expression: Version 6.0. Wageningen, the Netherlands: Noldus Information Technology. [Online]. Available: http://www.noldus.com/facereader

[30] P. Lewinski, T. M. den Uyl, and C. Butler, "Automated facial coding: Validation of basic emotions and facs aus in facereader," *J. Neurosci. Psychol. Econ.*, vol. 7, no. 4, pp. 227–236, 2014.

[31] H. van Kuilenburg, M. Wiering, and M. den Uyl, "A model based method for facial expression recognition," in *Proceedings of the 16th European Conference on Machine Learning*, 2005, pp. 194–205.

[32] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Conference on Computer Vision and Pattern Recognition*, 2001, pp. 511–518.

[33] J. Davis and A. F. Bobick, "The representation and recognition of human movements using temporal templates," in *Proceedings of the IEEE CVPR*, 1997, pp. 928–934.

[34] D. A. G. Jáuregui, L. Philip, C. Céline, S. Padovani, M. Bailly, and J.-C. Martin, "Video analysis of approach-avoidance behaviors of teenagers speaking with virtual agents," in *ICMI 2013*, 2013, p. 189196.

[35] T. Giraud, D. A. G. Jáuregui, J. Hua, B. Isableu, E. Filaire, C. L. Scanff, and J.-C. Martin, "Assessing postural control for affect recognition using video and force plates," in *ACII*, 2013, pp. 109–115.

[36] F. Cottin, C. Medigue, P. M. Lepretre, Y. Papelier, J. P. Koralsztein, and V. Billat, "Heart rate variability during exercise performed below and above ventilatory threshold," *Med. Sci. Sports Exerc.*, vol. 36, pp. 594–600, 2004.

[37] S. Sondhi, M. Khan, R. Vijay, and A. K. Salhan, "Vocal indicators of emotional stress," *Int. J. Comput. Appl.*, vol. 122, no. 15, pp. 38–43, 2015.