Progress in Neuropsychopharmacology & Biological Psychiatry 82 (2018) 107–114

Contents lists available at ScienceDirect

Progress in Neuropsychopharmacology & Biological Psychiatry

journal homepage: www.elsevier.com/locate/pnp

# Distinguish self- and hetero-perceived stress through behavioral imaging and physiological features

Michel Spodenkiewicz[a,c,d,*], Jonathan Aigrain[a], Nadège Bourvis[a,b], Séverine Dubuisson[a], Mohamed Chetouani[a], David Cohen[a,b,**]

[a] Sorbonne Universités, UPMC Univ Paris 06, CNRS UMR 7222, Institut des Systèmes Intelligents et de Robotique, Paris, France
[b] Département de Psychiatrie de l'Enfant et de l'Adolescent, AP-HP, GHU Pitié-Salpêtrière, Paris, France
[c] CEPOI EA 7388, Unité de Pédopsychiatrie de Liaison, Pôle de Santé Mentale, CHU Sud Réunion, Université de la Réunion, Saint-Pierre, France
[d] CESP Inserm U1178, Groupe IPSOM, Paris, France

A R T I C L E I N F O

*Keywords:*
Stress
Behavioral automatic assessment
Multimodality
Self-perception
Hetero-perception

A B S T R A C T

Stress reactivity is a complex phenomenon associated to multiple and multimodal expressions. Response to stressors has an obvious survival function and may be seen as an internal regulation to adapt to threat or danger. The intensity of this internal response can be assessed as the self-perception of the stress response. In species with social organization, this response also serves a communicative function, so-called hetero-perception. Our study presents multimodal stress detection assessment - a new methodology combining behavioral imaging and physiological monitoring for analyzing stress from these two perspectives. The system is based on automatic extraction of 39 behavioral (2D + 3D video recording) and 62 physiological (Nexus-10 recording) features during a socially evaluated mental arithmetic test. The analysis with machine learning techniques for automatic classification using Support Vector Machine (SVM) show that self-perception and hetero-perception of social stress are both close but different phenomena: self-perception was significantly correlated with hetero-perception but significantly differed from it. Also, assessing stress with SVM through multimodality gave excellent classification results (F1 score values: 0.9 ± 0.012 for hetero-perception and 0.87 ± 0.021 for self-perception). In the best selected feature subsets, we found some common behavioral and physiological features that allow classification of both self- and hetero-perceived stress. However, we also found the contributing features for automatic classifications had opposite distributions: self-perception classification was mainly based on physiological features and hetero-perception was mainly based on behavioral features.

## 1. Introduction

Stress can be approached by a wide range of scientific fields, depending on whether one refers to its definition (Koolhaas et al. 2011), its developmental and evolutionary function (Del Giudice et al., 2011), the contribution of genetic or environmental factors to explain species or individual variation (Laland et al. 2014; Wray et al., 2014), or the way it can be measured (Lutchyn et al. 2015). A narrow definition of stress describes these complex phenomena as "a condition where an environmental demand exceeds the natural regulatory capacities of an organism, in particular in situations that include unpredictability and uncontrollability" (Koolhaas et al. 2011). Despite the width of this field of research, some points seem consensual. (i) The stress response system participates to species survival and individual adaptation and implies immediate changes both on neurobiological and behavioral levels. (ii) Biological response relates to changes in the hypothalamo-pituitary-adrenal (HPA) pathway and the autonomic nervous system (ANS) that encode numerous short and long-term cascades (Szabo et al. 2012). (iii) Stress appears to have 3 main biological functions: it coordinates the organism's allostatic response to external and internal challenges both physical and psychosocial; it encodes, filters and reduces information about the organism's environment; it regulates, at both short and long term time scales, the physiology and behavior of a large range of social interaction areas (e.g. parenting, risk taking behavior in social context, coping behavior, reproduction, affiliation) (Del Giudice et al., 2011).

Functional convergence is a widespread phenomenon in evolution, revealing sometimes striking functional similarities between very distant species. In all species, stress is an internal mechanism to adapt to

* Correspondence to: M. Spodenkiewicz, Pôle de Santé Mentale, Centre Hospitalier Universitaire Sud Réunion, BP350, 97448 Saint-Pierre Cedex, La Réunion, France.
** Correspondence to: D. Cohen, Département de Psychiatrie de l'Enfant et de l'Adolescent, AP-HP, Hôpital Pitié-Salpêtrière, 47-83 Boulevard de l'Hôpital, 75651 Paris, Cedex 13, France.
*E-mail addresses:* michel.spodenkiewicz@chu-reunion.fr (M. Spodenkiewicz), david.cohen@aphp.fr (D. Cohen).

https://doi.org/10.1016/j.pnpbp.2017.11.023
Received 5 July 2017; Received in revised form 20 November 2017; Accepted 27 November 2017
Available online 27 November 2017

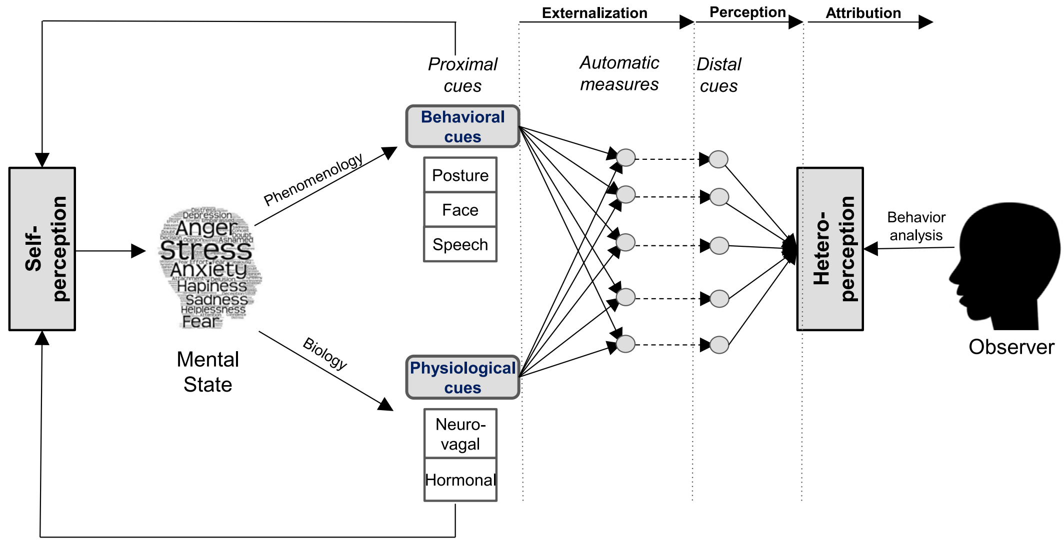


**Fig. 1.** Schematic illustration of the cascades following adaptation to stress and how behavioral and physiological changes may participate to both self-perception (left side) and hetero-perception (right side) of stress.

danger and may induce individual behavioral reaction. In animals with social interaction, reaction to stress is also associated with inter individual communication. For examples, distress from pups' early separation is revealed in rodents by using ultrasounds (Nagasawa et al. 2012). Alarm calls asking for protection or promoting protection of others exist in Velvet monkeys using different type of vocalization depending on the threat (Seyfarth et al. 1980). In humans at early stages of development, before the acquisition of mobility, infants seem to resort specifically to audio modality to inform their caregiver of their needs (e.g. infant manifests from separation by crying and caregivers seems to be specifically receptive to infant crying in hominins; Falk 2004). This “alarm call function” has been widely studied in child development literature (e.g. Soltis 2004; Weisman et al., 2015) and the parenting literature (e.g. Feldman 2015; Piallini et al. 2015). Therefore, from a systemic perspective, stress may be studied in humans, associating an internal view (self-perception) and an external/interactive view (hetero-perception, Fig. 1).

The key role of multimodality in affect detection is supported by a large body of literature in affective computing (see the survey of Sharma and Gedeon, 2012 and the review of D'Mello and Kory 2015). In terms of paradigms, the current methods support the emerging concept of behavior imaging and affective computing that aims at assessing affect and emotion through automatic and multimodal methods (Rehg et al. 2014; Leclère et al. 2016; Greene et al. 2016; Zhang et al. 2016). In psychiatry, applications have been developed for measuring depression (Joshi et al. 2013; Girard et al. 2014), anxiety (Hamilton 1959; Scherer et al. 2014), or autism (Rehg et al. 2014). Combined with method assessing interaction and synchrony (Delaherche et al. 2012; Leclère et al., 2014), applications have been proposed to assess patient-therapist during psychotherapies (Ramseyer and Tschacher 2014), infant-mother early interaction when one of the partner is dysfunctional be it the child (e.g. infants developing autism) (Saint-Georges et al. 2011; Cohen et al. 2013) or the mother (e.g. mothers showing severe psychopathology) (Hammal et al. 2015; Leclère et al., 2014).

From the many physiological and behavioral changes triggered by a stressful experience, it is still difficult to understand those that participate to internal experience of stress from those that contributes to its communicative dimension (Aigrain et al. 2016). Given the multimodal characteristics of short-term response to social stress, our hypothesis is that changes in the ANS mainly contribute to the self-perception, whereas behavioral changes mainly contribute to the hetero-perception. In this experiment, we propose a novel method to assess multimodal changes associated to acute social stress response simultaneously from both self- and hetero-perspectives.

## 2. Material and methods

Twenty-five individuals (mean age = 26.3 ± 4.6 years, 64% female) participated to a socially evaluated mental arithmetic test. The test was composed of 6 steps of increasing difficulty with a break period of 5 s between 2 steps. Principals were inspired from a mental arithmetic task used for the validation of the Mathematical Anxiety Rating Scale (Ashcraft and Faust, 1994). We asked participants a succession of computational problems (e.g. addition of time values) of increasing complexity. The participants were told that the aim of the experiment was to assess their cognitive performance, and that both quickness and correctness of answers were taken into account to compute the score. A false score bar was projected on the screen showing that the subject's performance was below average. Once the test was finished, we revealed that triggering stress was the real purpose of the experiment.

All participants were recruited among medical students of the *Université Pierre et Marie Curie* in Paris, after oral and written informed consent. Because of acquisition problems, we dismissed data of 4 participants. The experimental setup and the feature acquisition are described in Fig. 2. The experimental room had a screen to show the task, a Kinect sensor to record video data from the whole body, and an HD camera to record the video of facial movements. Physiological sensors (Nexus-10) were used to record Heart Rate, Respiratory Rate, Skin Conductance and Temperature.

Self-assessment of stress was conducted immediately after the task. The subjects answered a Likert-scaled (1–5) question about how stressed they felt during each step. They had the possibility to watch their own videos before providing their answers to limit memory bias such as fading affect bias or misattribution of memory. Indeed, both anxiety (threat-of-shock or ego-threat) and mathematical task reliably affect working memory performance. They cause deficits in attentional control and impair the ability to inhibit irrelevant information or maintain the relevant one (Moran 2016). Then, in order to obtain binary labels, we use a threshold on the stress level: Non-Stress = {1, 2} and Stress = {3, 4, 5}. The threshold was based on previous analysis (see Aigrain et al. 2016).

Hetero-assessment was assessed using the crowdsourcing platform *CrowdFlower* (www.crowdflower.com). We uploaded and presented the videos of the participants during the experiment to adults' annotators from various countries, only telling them that the videotaped subjects

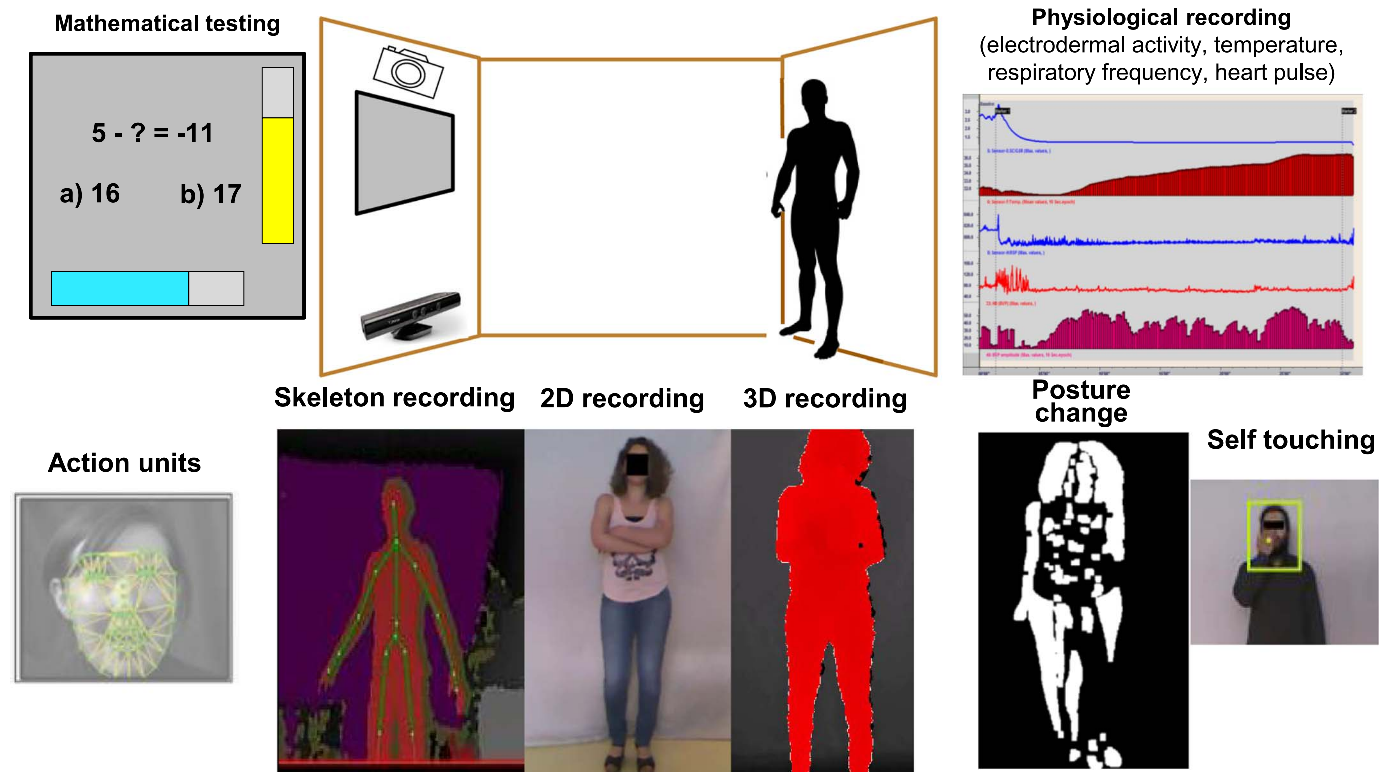


**Fig. 2.** Experimental setup and feature acquisition.
Experimental set up (up-middle) and arithmetic task (up-left); feature acquisition: physiological variables (up-right), 2D, 3D and skeleton recording (down-middle), action units (down-left) and behavioral extraction (e.g. self-touching, down-right).

were taking a cognitive test. Three questions were asked online after each video: *Do you think this person is stressed*? (Answers: not stressed/stressed); *How stressed is the person in this video*? (Answers: Likert scale 1–5); *How confident are you on your ratings*? (Answers: Likert scale 1–5). 10 annotations per subject were done. We used 3 mechanisms to ensure the annotation quality: (1) highest ranked category annotators; (2) minimum time setting to rate the videos; (3) test question for credibility. 248 people annotated an average of 6.45 ± 5.22 videos. After removing untrustworthy annotators using answers to Test Questions, we used the Honeypot method (Nguyen et al., 2013) to assign a single label to each video,' (if more than half of the remaining annotations were Stress answers, we assigned the Stress label, otherwise we assigned the Non-Stress label). Finally, we extracted the answers as data for further analysis.

From the recordings, we extracted automatically 39 behavioral and 62 physiological features presented in Table 1. (i) From the high definition video recording of the face, we extracted 12 Action Units (AU) using Nicolle et al. (2015) methods. (e.g. AU1 = inner brow raiser; AU6 = cheek raiser; AU25 = lips part). AUs are part of the Facial Action Coding System (Ekman and Friesen, 1978; Hamm et al. 2011), one of the standards for systematic categorization of facial expressions. (ii) From the RGB video recordings, we extracted 7 Quantity of Movement features (QoM) (the number of pixels that changed between two successive frames) and the sum of the displacements of the skeleton joints for the whole body as well as head and hands. For these features, normalization was computed according to the size of the participants and the distance from the camera during the experiment (Aigrain et al. 2016). (iii) From RGB video recording, we monitored 8 features from automatic detection of periods of high body activity, posture changes and face self-touching using a method detailed in Aigrain et al. 2015. (iv) From the physiological recordings, we extracted cardiac functions (blood volume pulse, heart rate and heart rate variability), respiratory system parameters (chest and abdominal respiration), electrodermal activity, skin temperature and electromyographic activity of the sternocleidomastoid and upper trapezius (Sharma and Gedeon 2012). We used as features the mean, the variance, the minimum and the maximum values of the recorded or extracted signals as it has been done in previous studies implying human-computer interaction to recognize emotions from physiological signals (Lisetti and Nasoz 2004).

To explore how behavioral and physiological features participated to either self- or hetero-perception of stress, we used machine-learning methods to select the best classification feature set for each perception. The different steps are detailed in Fig. 3. First, we used the Box−Cox transformation, to normalize the feature distributions (Sakia 1992). Second, we performed feature subset selection in order to avoid overfitting and better understand the predictive power of each feature. We tested 3 different methods (details are given in Aigrain et al. 2016): (i) Forward selection wrapper (FSW): Wrappers evaluate a subset of features by using the same machine learning algorithm as in the final application (Kohavi and John 1997). In our case, we use a support machine vector (SVM) with a linear kernel function. Since training SVMs is computationally expensive, exploring the space of feature subsets is usually done using greedy methods (Guyon and Elisseeff 2003). With forward selection, starting from using only the feature with the best accuracy, we iteratively add the best feature among the remaining ones. Once all features have been added, we keep the subset that gives the best classification performances. (ii) Backward elimination wrapper (BEW): This method also uses a SVM to evaluate subsets. Starting from the complete set of features, we iteratively remove the worst feature of the remaining set. Once all features have been removed, we keep the subset that gives the best classification performances. (iii) Simulated annealing with Hall correlation (SAHC): For this method, we use the simulated annealing met a heuristic (Kirkpatrick et al. 1983) to explore the space of feature subsets. Because of the computational cost of this space search strategy, we use the Hall correlation to evaluate feature subsets. We then get a good approximation

**Table 1**
List of the behavioral and physiological features extracted during the stress experiment.

| Main domains | Features | References |
|---|---|---|
| Behavioral features | | |
| Face action unit (mean, SD) | AU1: Inner Brow Raiser; AU2: Outer Brow Raiser; AU4: Brow Lowerer; AU5: Upper Lid Raiser; AU6: Cheek Raiser; AU9: Nose Wrinkler; AU12: Lip Corner Puller; AU15: Lip Corner Depressor; AU17: Chin Raiser; AU20: Lip Stretcher; AU25: Lips Part; AU26: Jaw Drop | Nicolle et al. 2015 |
| Quantity of movement (mean) | QoM computed with the skeleton; QoM computed with the RGB frames; QoM for the left hand; QoM for the right hand; QoM for both hands; QoM for the head; QoM for the head only along Z-axis | Aigrain et al. 2016 |
| Specific body features | Number of periods of high activity; Mean duration of periods of high activity; Mean highest value of periods of high activity; Number of posture changes; Number of times face touching with one hand occurred; Mean duration of face touching with one hand; Number of times face touching with two hands occurred; Mean duration of face touching with two hands | Aigrain et al. 2015 |
| Physiological features | | |
| Blood volume pulse (mean, var., min, max) | Blood volume pulse; Blood volume pulse amplitude | Barreto et al., 2007 |
| Electromyographic activity (mean, var., min, max) | Electromyographic activity of the sternocleidomastoid and upper trapezius - channel 1; Electromyographic activity of the sternocleidomastoid and upper trapezius - channel 2; Electromyographic activity of the sternocleidomastoid and upper trapezius Mean Frequency; Electromyographic activity of the sternocleidomastoid and upper trapezius Amplitude | Wijsman et al. 2011 |
| Skin (mean, var., min, max) | Electrodermal activity; Temperature | Sharma and Gedeon 2012 |
| Heart rate (mean, var., min, max) | Heart Rate; Heart Rate Variability Amplitude; Heart Rate Variability Low Frequency zone; Heart Rate Variability square root of the mean squared difference between adjacent N—N interval*; Heart Rate Variability Standard Deviation of Normal to Normal intervals* | Sharma and Gedeon 2012 |
| Respiration (mean, var., min, max) | Chest and abdominal Respiration; Chest and abdominal Respiration Amplitude; Chest and abdominal Respiration Rate | Healey and Picard 2005 |
| Heart rate and respiration (mean, var., min, max) | Level of coherence between the Respiration and the Heart Rate | Sharma and Gedeon 2012 |

QoM: Quantity of movement; AU: Action unit; var.: variance; min: minimum; max: maximum.*one metric only.

of the subset that both maximizes the correlation between features and labels and minimizes the inter-feature correlation.

Finally, we used SVMs with linear kernel function to process a classification task (Stress or Non-Stress) and to evaluate the predictive value of our framework for each configuration of perception (self vs. hetero perception). In a previous technical paper, we showed that SVM with linear kernel function performed better than SVM with either polynomial or radial basis functions (Aigrain et al. 2016). We used a 10-fold subject independent cross validation strategy to compute the results: steps from 2 or 3 people were used as the testing set. The steps of the remaining people are used as the training set. This cross validation was also used with the training set to determine the SVM and kernel function parameters.

We chose the F1 score for both Stress and Non-Stress classes as the performance metric, since our dataset is unbalanced for self-perception and for hetero-perception. We used the Student's *t*-test to compare two F1 score values. He we present: F1 scores that display the results obtained by each feature when used alone for each assessment set; and F1 scores that display the results of the best subsets of features for each assessment. The baseline F1 score (chance level) shown in Fig. 4 was

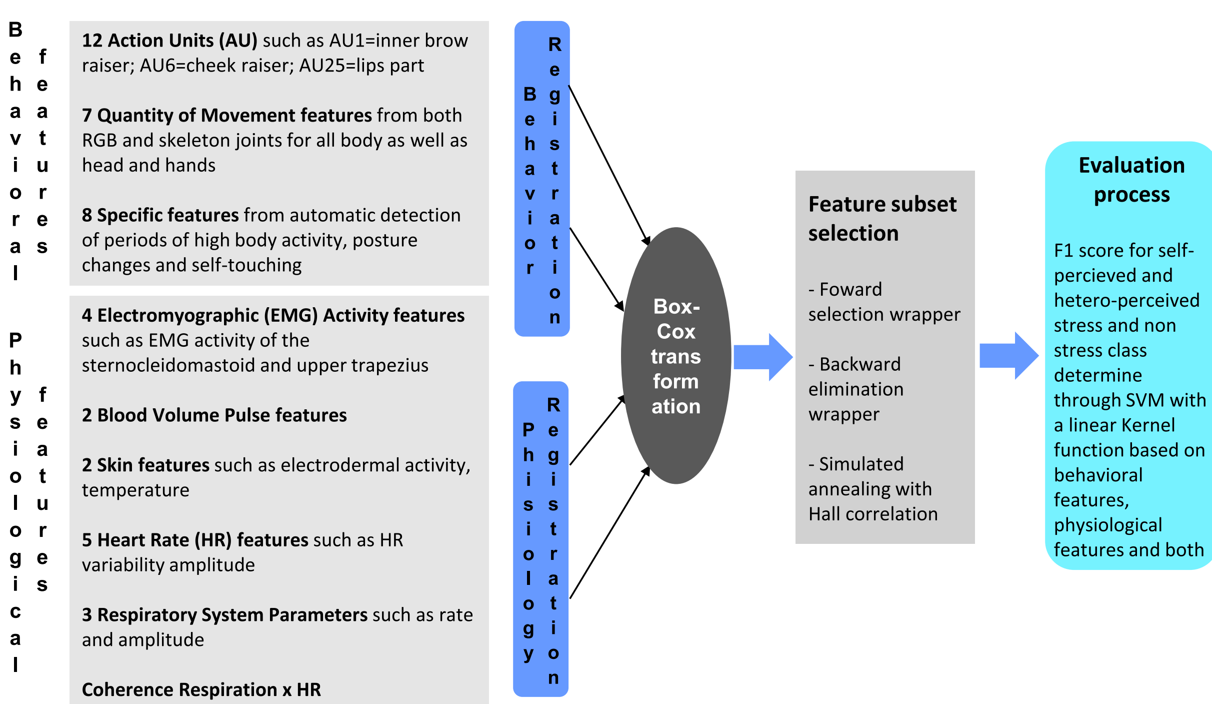


**Fig. 3.** Diagram flow of the different steps to select best classification feature sets.

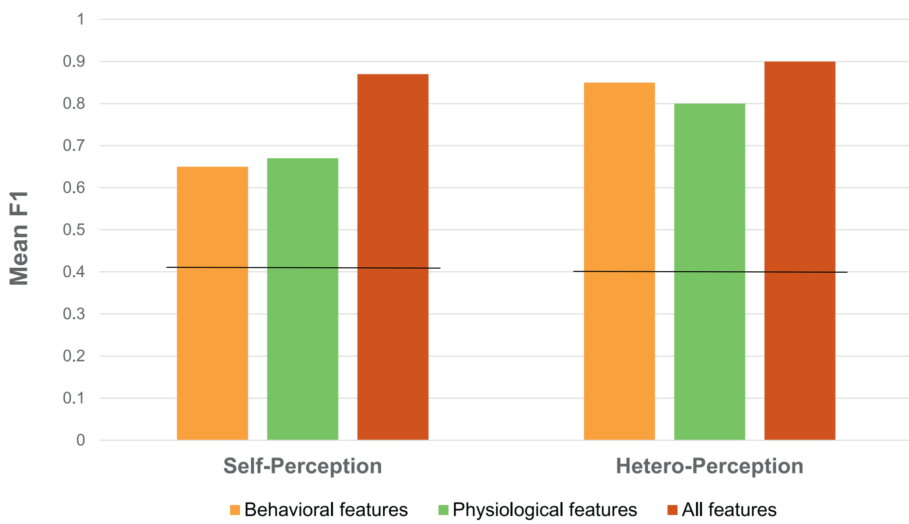


**Fig. 4.** Prediction performances of each SVM classification for self-perception and hetero-perception.
The baseline average F1 score obtained by a random classifier for self-perception is 0.404 ( ± 0.079). The baseline average F1 score obtained by a random classifier for hetero-perception is 0.41 ( ± 0.083). The chance levels are indicated by the grey lines on the figure.

obtained by a random classifier.

## 3. Results

We first examined how the participants' videos were perceived in self vs. hetero assessments. There was a significant difference with a higher rate of stress-rated videos in the self-perception assessment (stress = 74.6% and non-stress = 25.4% in self-perception vs. stress = 60.3% and non-stress = 39.7% in hetero-perception, $p < 0.05$). To explore the relationship between self-perception and hetero-perception, we calculated Cohen's Kappa based on binary labels and correlation coefficient based on non-binary values. We found a significant moderate association ($\kappa = 0.38$ and $\rho = 0.41$, $p < 0.05$) indicating that the two phenomena are correlated but not similar. Combining the two results means that individuals rating themselves tended to experience stress earlier during the socially evaluated mental arithmetic test than what was perceived by an external observer.

Second, we explored behavioral and physiological features that were associated with either self-perceived or hetero-perceived stress. Fig. 4 shows the SVM classification results obtained by the best selected feature subsets for each assessment. Features were selected from either (i) the whole set of features (brown), (ii) only behavioral ones (orange) or (iii) only physiological ones (green). Regarding self-perception (Fig. 4, left), we found that the combination of physiological and behavioral features outperformed the results obtained when using only one modality. Since the participants of the experiment watched their own videos before annotating them, their answers were the result of both their personal experiences and their behavior analysis. Regarding hetero-perception (Fig. 4, right), we found that both modalities achieve good F1 scores: 0.85 ( ± 0.020) for behavioral features and 0.8 ( ± 0.021) for physiological ones. It is not surprising that behavioral features significantly outperform physiological ones ($p < 0.0001$) since annotators based their judgment by watching the behavior of another person in each video.

Third, we looked for common and specific physiological and behavioral features found in the best selected feature subsets. These are listed in Table 2 for self-perception (left) and hetero-perception (right). For self-perception, the best subset is composed of 32 features: 21 physiological ones and 11 behavioral ones. For hetero-perception, the best subset is composed of 24 features: 9 physiological ones and 15 behavioral ones. Looking at common features that contributed to both classifications, we found 5 behavioral features (3 AU related to smile and brow lowerer: Inner and Out Brow Raisers, Lip Stretcher and Lip Corner Depressor, Chin Raiser, and Jaw Drop; 2 related to QoM) and 6 physiological ones (3 related to respiration, 1 to electrodermal activity, 1 to Blood Volume Pulse, and 1 to EMG). Regarding specific features, we found a significant opposite distribution of behavioral and physiological features between self-perception and hetero-perception. There were 5 behavioral features (1 related to Periods of High Activity of the whole body; 2 AU related to facial movements: Chin and Cheek Raisers – SD; 2 related to Face Touching with one or both hands – count) and 16 physiological ones (4 related to EMG activity of the sternocleidomastoid & upper trapezius; 1 to maximal Heart Rate; 1 related to Heart Rate Variability Amplitude; 4 related to Blood Volume Pulse; 5 to Chest and abdominal Respiration; 1 related to Temperature) for the self-perception best subset (Chi2 = 7.81, df = 1, p = 0.005), as opposed to 10 behavioral features (1 related to Head Movements; 1 related to Posture Change Count; 7 AU related to facial movements: Inner and Outer Brow Raisers, Chin Raiser – mean, Lip Stretcher, Jaw Drop; 1 related to Face Touching with one Hand Duration) and 3 physiological ones for the hetero-perception (1 related to Blood Volume Pulse – mean; 2 related to Levels of coherence between the Respiration and the Heart Rate) best subset (Chi2 = 6.79, df = 1, p = 0.009).

## 4. Discussion

We present a new and original method to capture multimodal assessment of social stress. The current experiment is based on (1) the multimodal expression of stress reactivity, and (2) the evolutionary perspective that confers to stress reactivity both a survival and a communicative function. The results support this view by showing that self-perception and hetero-perception are indeed both close and different phenomena. We found that self-perception was correlated with hetero-perception but qualitatively differed from it. Also, assessing stress with machine learning methods through multimodality gave excellent classification results for both self-perception and hetero-perception.

In the best selected feature subsets, we found some common behavioral and physiological features participating to classification of both self-perception and hetero-perception. However, we also found many specific features with an opposite distribution of behavioral and physiological features between self-perception (more physiological features) and hetero-perception (more behavioral features). In this experiment, the physiological features specific for self-perception of stress include several metrics associated to Blood Volume Pulse (Barreto et al., 2007), Heart Rate Variability (Aigrain et al. 2016; Acerbi et al. 2017), Chest and Abdominal Respiration (Healey and Picard 2005; Plarre et al. 2011) and EMG activity (Healey and Picard 2005; Wijsman et al. 2011). The ranking of features for measuring self-perceived stress is coherent with the current state of art: the best scores are obtained from signals

**Table 2**
Behavioral and physiological characteristics of self-perceived and hetero-perceived stress: common and specific features detected through machine learning methods.

| | F1 | Self-perception | Hetero-perception | F1 |
|---|---|---|---|---|
| | Common features | | | |
| Behavioral features | 0.62 | Quantity of Movement | | 0.73 |
| | 0.60 | Right Hand Movement | | 0.67 |
| | 0.52 | Face AU 6 (Cheek Raiser): mean | | 0.61 |
| | 0.48 | Face AU 12 (Lip Corner Puller): SD | | 0.62 |
| | 0.43 | Face AU 4 (Brow Lowerer): mean | | 0.46 |
| Physiological features | 0.56 | Level of coherence between the Respiration and the Heart Rate: max | | 0.56 |
| | 0.51 | Blood Volume Pulse Amplitude: max | | 0.69 |
| | 0.51 | Chest and Abdominal Respiration Rate: maxi | | 0.53 |
| | 0.50 | Chest and Abdominal Respiration: var | | 0.64 |
| | 0.50 | Galvanic Skin Response: var | | 0.48 |
| | 0.43 | EMG activity of the sternocleidomastoid & upper trapezius channel 1: min | | 0.49 |
| | Specific features | | | |
| Behavioral features | 0.62 | Mean (highest value) of Periods of High Activity | Head Movement | 0.78 |
| | 0.57 | Face AU 6 (Cheek Raiser): SD | Posture Change Count | 0.60 |
| | 0.51 | Face Touching with one Hand Count | Face AU 17 (Chin Raiser): mean | 0.60 |
| | 0.50 | Face AU 17 (Chin Raiser): SD | Face AU 20 (Lip Stretcher): SD | 0.59 |
| | 0.34 | Face Touching with two Hands Count | Face AU 1 (Inner Brow Raiser): SD | 0.58 |
| | | | Face Touching with one Hand Duration | 0.57 |
| | | | Face AU 26 (Jaw Drop): SD | 0.55 |
| | | | Face AU 15 (Lip Corner Depressor): mean | 0.53 |
| | | | Face AU 2 (Outer Brow Raiser): mean | 0.52 |
| | | | Face AU 2 (Outer Brow Raiser): SD | 0.52 |
| Physiological features | 0.56 | EMG activity of the sternocleidomastoid & upper trapezius channel 2: var | Blood Volume Pulse: mean | 0.67 |
| | 0.55 | Heart Rate: max | Level of coherence between the Respiration and the Heart Rate: mean | 0.56 |
| | 0.55 | Heart Rate Variability Amplitude: var | Level of coherence between the Respiration and the Heart Rate: min | 0.51 |
| | 0.54 | Blood Volume Pulse: min | | |
| | 0.52 | Chest & abdominal Respiration Rate: min | | |
| | 0.51 | Blood Volume Pulse Amplitude: var | | |
| | 0.50 | Temperature: min | | |
| | 0.49 | EMG activity of the sternocleidomastoid & upper trapezius channel 2: mean | | |
| | 0.49 | Chest & abdominal Respiration Amplitude: var | | |
| | 0.48 | EMG activity of the sternocleidomastoid & upper trapezius Mean Frequency: var | | |
| | 0.47 | Chest & abdominal Respiration Amplitude: max | | |
| | 0.44 | Blood Volume Pulse Amplitude: min | | |
| | 0.44 | Chest & abdominal Respiration Rate: mean | | |
| | 0.42 | Chest & abdominal Respiration Amplitude: min | | |
| | 0.41 | Blood Volume Pulse: max | | |
| | 0.35 | EMG activity of the sternocleidomastoid & upper trapezius Mean Frequency: max | | |

SD: standard deviation; min: minimum; max: maximum; var.: variance; AU: Action Unit; EMG: electromyographic.

linked with the cardiac activity and the less accurate ones with EMG (Sharma and Gedeon 2012; Acerbi et al. 2017). The behavioral features specific for hetero-perception of stress include several metrics associated to facial AUs (related to mouth and eyebrow movement) that are mainly associated to negative emotions such as sadness and disgust (Ekman et al. 1980). They have been previously associated with high- and low- stressor performance (Dinges et al. 2005). Behavioral features also include head movement, face touching, and posture change (Harrigan 1985; Giannakakis et al. 2017).

Since the current method may help disentangling self-perceived reaction to stress and hetero-perceived reaction to stress by offering differential multimodal metrics, we speculate that patients with stress and emotional dysregulation could be investigated with behavioral imaging. In patients suffering from borderline personality disorder with hyperarousal to stressors (fast behavioral response to soft social stimuli) (Links et al., 2017) and poor insight (including poor self-perception of their own emotional internal state) (Zanarini and Frankenburg 2007; Damman et al. 2011; Spodenkiewicz et al. 2013), we hypothesize a dissociation between objective automatic measures of stress (hetero-perceived stress) and subjective (self-perceived) stress.

The results should be discussed within the context of the study strengths and limitations. The strengths include (i) the multi-perspective and multimodal approach of the stress phenomenon with a careful selection of potential predictive features (Aigrain et al. 2016); (ii) the introduction of new behavioral features for stress detection such as face-touching, (iii) the high number of extracted features and (iv) the high classification scores that allow us to interpret the predictive power of some features. The limitations include (i) the experimental stimulus that elicits stress in a specific context that also includes cognitive load to social stress; (ii) the sample size that limited statistical power. Thus, for statistical reasons we considered stress as a discrete variable. Further studies based on more participants will have to address the issue with stress analyzed as a continuous parameter. (iv) Finally, even if we adopted a multimodal perspective and monitored 101 features, we did not address important behavioral features of stress convey through speech (Zhou et al. 2001), or physiological ones convey through hormonal pathways (Gordis et al. 2006; Nater et al. 2005; Takai et al. 2004). Also, future studies should compare differences between genders considering previous studies (Wang et al. 2007, Streit et al., 2017).

## 5. Conclusions

Self-perception and hetero-perception of stress are both close but different phenomena. New methods of automatic assessment of social stress (self-perceived or hetero-perceived stress) with multimodal techniques combining behavior imaging and physiological monitoring may give excellent classification results. In the best selected feature subsets found in this experiment, we had some common behavioral and physiological features participating to classification of both self-perception and hetero-perception. However, we also found specific features with an opposite distribution of behavioral and physiological features between self-perception (mainly physiological features) and hetero-perception (mainly behavioral features). It is likely that, during the process of evolution, response to stress has selected from a common set of behavioral and physiological features some to support internal perception and others to support inter-individual communication.

## Role of the funding source

This work was partially supported by the Labex SMART (ANR-11-LABX-65) under French state funds managed by the ANR within the Investissements d'Avenir programme (ANR-11-IDEX-0004-02); by the Fondation pour la Recherche Médicale (DPP20151033945); and by the Groupement de Recherche en Psychiatrie (GDR-3557). Sponsors had no involvement in study design, data analysis, or interpretation of results.

## Acknowledgements

The research was approved by the Institutional Review Board at University Pierre et Marie Curie, was conducted according to ethical standards, and all participating signed an informed consent. All co-authors.

agree with the contents of the manuscript and there is no financial interest to report.

## References

Acerbi, G., Rovini, E., Betti, S., Tirri, A., Rónai, J.F., Sirianni, A., Agrimi, J., Eusebi, L., Cavallo, F., 2017. A wearable system for stress detection through physiological data analysis. Ital. Forum of Ambient Assist. Living, Lect. Notes 31–50.

Aigrain, J., Dubuisson, S., Detyniecki, M., Chetouani, M., 2015. Person-specific behavioural features for automatic stress detection. IEEE Int. Conf. Workshops on Autom. Face and Gesture Recogn. http://dx.doi.org/10.1109/FG.2015.7284844.

Aigrain, J., Spodenkiewicz, M., Dubuisson, S., Detyniecki, M., Cohen, D., Chetouani, M., 2016. Multimodal stress detection from multiple assessments. IEEE Trans. on Affect. Comput. (in press). http://dx.doi.org/10.1109/TAFFC.2016.2631594.

Ashcraft, M., Faust, M.W., 1994. Mathematics anxiety and mental arithmetic performance: an exploratory investigation. Cogn. Emotion 8 (2).

Barreto, A., Zhai, J., Adjouadi, M., 2007. Non-intrusive physiological monitoring for automated stress detection in human-computer interaction. Hum.-Comput. Interact. 29–38.

Cohen, D., Cassel, R.S., Saint-Georges, C., Mahdhaoui, A., Laznik, M.C., Apicella, F., Muratori, P., Maestro, S., Muratori, F., Chetouani, M., 2013. Do parentese prosody and Fathers' involvement in interacting facilitate social interaction in infants who later develop autism? PLoS One 8 (5), e61402. http://dx.doi.org/10.1371/journal.pone.0061402.

Damman, G., Hugli, C., Selinger, J., Gremaud-Heitz, D., Sollberger, D., Wiesbck, G.A., Kuchenhoff, J., Walter, M., 2011. The self-image in borderline personality disorder: an in-depth qualitative research study. J. Personal. Disord. 25 (4), 517–527.

Del Giudice, M., Ellis, B.J., Shirtcliff, E.A., 2011. The Adaptive Calibration Model of stress responsivity. Neurosci. Biobehav. Rev. 35 (7) (15, 62-92).

Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., Cohen, D., 2012. Interpersonal synchrony: a survey of evaluation methods across disciplines. IEEE Trans. Affect. Comput. 3 (3), 349365.

Dinges, D.F., Rider, R.L., Dorrian, J., McGlinchey, E.L., Rogers, N.L., Cizman, Z., Goldenstein, S.K., Vogler, C., Venkataraman, S., Metaxas, M.N., 2005. Optical computer recognition of facial expressions associated with stress induced by performance demands. Aviat. Space Environ. Med. 76, 172–182.

D'Mello, S.K., Kory, J., 2015. A review and meta-analysis of multimodal affect detection systems. ACM Comput. Surv. 47 (3) (43, 1–36). https://doi.org/10.1145/2682899.

Ekman, P., Friesen, W., 1978. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto.

Ekman, P., Freisen, W.V., Ancoli, S., 1980. Facial signs of emotional experience. J. Pers. Soc. Psychol. 39 (6), 1125–1134. http://dx.doi.org/10.1037/h0077722.

Falk, D., 2004. Prelinguistic evolution in early hominins: whence motherese? Behav. Brain Sci. 27 (4), 491–503.

Feldman, R., 2015. The neurobiology of mammalian parenting and the biosocial context of human caregiving. Horm. Behav. 77, 3–17. http://dx.doi.org/10.1016/j.yhbeh.2015.10.001.

Giannakakis, G., Pediaditis, M., Manousos, D., Kazantzaki, E., Chiarugi, F., Simos, P., Marias, K., Tsiknakis, M., 2017. Stress and anxiety detection using facial cues from videos. Biomed. Signal Proc. and Control 31, 89–101. http://dx.doi.org/10.1016/j.bspc.2016.06.020.

Girard, J.M., Cohn, J.F., Mahoor, M.H., Mavadati, S.M., Hammal, Z., Rosenwald, D.P., 2014. Nonverbal social withdrawal in depression: evidence from manual and automatic analyses. Image Vis. Comput. 32 (10), 641–647. http://dx.doi.org/10.1016/j.imavis.2013.12.007.

Gordis, E., Granger, D., Susman, E., Trickett, P., 2006. Asymmetry between salivary cortisol and α-amylase reactivity to stress: relation to aggressive behavior in adolescents. Psychoneuroendocrinology 31 (8), 976–987. http://dx.doi.org/10.1016/j.psyneuen.2006.05.010.

Greene, S., Thapliyal, H., Caban-Holt, A., 2016. A survey of affective computing for stress detection: evaluating technologies in stress detection for better health. IEEE Consum. Electron. Mag. 5, 44–56. http://dx.doi.org/10.1109/MCE.2016.2590178.

Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. J. Mach. Learn. Res. 3, 1157–1182.

Hamilton, M., 1959. The assessment of anxiety-states by rating. Br. J. Med. Psychol. 32, 50–55.

Hamm, J., Kohler, C.G., Gur, R.C., Verma, R., 2011. Automated facial action coding system for dynamic analysis of facial expressions in neuropsychiatric disorders. J. Neurosci. Methods 200 (2), 237–256. http://dx.doi.org/10.1016/j.jneumeth.2011.06.023.

Hammal, Z., Cohn, J.F., Messinger, D.S., 2015. Head movement dynamics during play and perturbed mother-infant interaction. IEEE Trans. Affect. Comput. 6 (4), 361370.

Harrigan, J.A., 1985. Self-touching as an indicator of underlying affect and language processes. Soc. Sci. Med. 20 (11), 1161–1168. http://dx.doi.org/10.1016/0277-9536(85)90193-5.

Healey, J.A., Picard, R.W., 2005. Detecting stress during real-world driving tasks using physiological sensors. IEEE Trans. Intell. Transp. Syst. 6 (2), 156–166. http://dx.doi.org/10.1109/TITS.2005.848368.

Joshi, J., Goecke, R., Parker, G., Breakspear, M., 2013. Automatic Face and Gesture Recognition 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition. http://dx.doi.org/10.1109/FG.2013.6553796.

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220 (4598), 671–680. http://dx.doi.org/10.1126/science.220.4598.671.

Kohavi, R., John, G.H., 1997. Wrappers for feature subset selection. Artif. Intell. 97, 273–324.

Koolhaas, J.M., Bartolomucci, A., Buwalda, B., De Boer, S.F., Flügge, S., Korte, S.M., Meerlo, P., Murison, R., Olivier, B., Palanza, P., Richter-Levin, G., Sgoifo, A., Steimer, T., Stiedl, O., Van Dijk, G., Wohr, M., Fuchs, E., 2011. Stress revisited: a critical evaluation of the stress concept. Neurosci. Biobehav. Rev. 35 (5), 291–301.

Laland, K., Uller, T., Feldman, M., Sterelny, K., Müller, G.B., Moczek, A., Jablonka, E., Odling-Smee, J., Wray, G.A., Hoekstra, H.E., Futuyma, D.J., Lenski, R.E., Mackay, T.F., Schluter, D., Strassmann, J.E., 2014. Does evolutionary theory need a rethink? Nature 514 (7521), 161–164.

Leclère, C., Viaux, S., Avril, M., Achard, C., Chetouani, M., Missonnier, S., Cohen, D., 2014. Why synchrony matters during mother-child interactions: a systematic review. PLoS One 9 (12), e113571. doi https://doi.org/10.1371/journal.pone.0113571.

Leclère, C., Avril, M., Viaux-Savelon, S., Bodeau, N., Achard, C., Missonnier, S., Keren, M., Feldman, R., Chetouani, M., Cohen, D., 2016. Interaction and behaviour imaging: a novel method to measure mother-infant interaction using video 3D reconstruction. Transl. Psychiatry 6 (5), e816. http://dx.doi.org/10.1038/tp.2016.82.

Links, P.S., Eynan, R., Heisel, M.J., Barr, A., Korzekwa, M., McMain, A., Ball, J.S., 2017. Affective instability and suicidal ideation and behavior in patients with borderline personality disorder. J. Personal. Disord. 21 (1), 72–86. http://dx.doi.org/10.1521/pedi.2017.21.1.72.

Lisetti, C.L., Nasoz, F., 2004. Using noninvasive wearable computers to recognize human emotions from physiological signals. J. App. Signal Proc. (11), 1672–1687.

Lutchyn, Y., Mark, G., Sano, A., Johns, P., Czerwinski, P., Iqbal, S., 2015. Stress is in the eye of the beholder. Int. Conf. Affect. Comput. Intell. Interact. http://dx.doi.org/10.1109/ACII.2015.7344560.

Moran, T., 2016. Anxiety and working memory capacity: a meta-analysis and narrative review. Psychol. Bull. 142. http://dx.doi.org/10.1037/bul0000051.

Nagasawa, M., Okabe, S., Mogi, K., Kikusui, T., 2012. Oxytocin and mutual communication in mother-infant bonding. Front. Hum. Neurosci. 6 (31). http://dx.doi.org/10.3389/fnhum.2012.00031.

Nater, U.M., Rohleder, N., Gaab, J., Berger, S., Jud, A., Kirschbaum, C., Ehlert, U., 2005. Human salivary alpha-amylase reactivity in a psychosocial stress paradigm. Int. J. Psychophysiol. 55 (3), 333–342. http://dx.doi.org/10.1016/j.ijpsycho.2004.09.009.

Nguyen, Q.V.H., Nguyen, T.T., Lam, Ngoc T., Aberer, K., 2013. An evaluation of aggregation techniques in crowdsourcing. In: Web Information Systems Engineering. Springer Berlin Heidelberg, pp. 1–15.

Nicolle, J., Bailly, K., Chetouani, M., 2015. Facial action unit intensity prediction via hard multi-task metric learning for kernel regression. IEEE Int. Conf. Workshops on Automatic Face and Gesture Recogn. http://dx.doi.org/10.1109/FG.2015.7284868.

Piallini, G., De Palo, F., Simonelli, A., 2015. Parental brain: cerebral areas activated by infant cries and faces. A comparison between different populations of parents and not. Front. Psychol. 6, 1625.

Plarre, K., Raij, A., Hossain, S.M., Ali, A.A., 2011. Continuous inference of psychological stress from sensory measurements collected in the natural environment. In: Conference: Proceedings of the 10th International Conference on Information Processing in Sensor Networks, IPSN. 2011.

Ramseyer, F., Tschacher, W., 2014. Nonverbal synchrony of head- and body-movement in psychotherapy: different signals have different associations with outcome. Front. Psychol. 5 (979). http://dx.doi.org/10.3389/fpsyg.2014.00979.

Rehg, J.M., Rozga, A., Abowd, G.D., 2014. Behavioral imaging and autism. IEEE Pervasive Comput. 13 (2), 84–87. http://dx.doi.org/10.1109/MPRV.2014.23.

Saint-Georges, C., Mahdhaoui, A., Chetouani, M., Cassel, R.S., Laznik, M.C., Apicella, F., Muratori, P., Maestro, S., Muratori, F0, Cohen, D., 2011. Do parents recognize autistic deviant behavior long before diagnosis? Taking into account interaction using computational methods. PLoS One 6 (7), e22393. http://dx.doi.org/10.1371/journal.pone.0022393.

Sakia, R.M., 1992. The box-cox transformation technique: a review. J. R. Stat. Soc. 41 (2), 169–178.

Scherer, S., Stratou, G., Lucas, G., Mahmoud, M., Boberg, J., Gratch, J., 2014. Automatic audiovisual behavior descriptors for psychological disorder analysis. Image Vis. Comput. 32 (10), 648–658. http://dx.doi.org/10.1016/j.imavis.2014.06.001.

Seyfarth, R.M., Cheney, D.L., Marler, P., 1980. Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. Science 201, 801–803.

Sharma, N., Gedeon, T., 2012. Objective measures, sensors and computational techniques for stress recognition and classification: a survey. Comput. Methods Prog. Biomed. 108, 1287–1301. http://dx.doi.org/10.1016/j.cmpb.2012.07.003.

Soltis, J., 2004. The signal functions of early infant crying. Behav. Brain Sci. 27 (4), 443458.

Spodenkiewicz, M., Speranza, M., TaÏeb, O., Pham-Scottez, A., Corcos, M., Révah-Levy, A., 2013. Living from day to day – qualitative study on borderline personality disorder in adolescence. J. Canad. Acad. Child Adolesc. Psychiatry 22 (4), 282–289.

Streit, F., Akdeniz, C., Haddad, L., Kumsta, R., Entringer, S., Frank, J., 2017. Sex-specific association between functional neuropeptide S receptor gene (NPSR1) variants and cortisol and central stress responses. Psychoneuroendocrinology 76, 49–56. http://dx.doi.org/10.1016/j.psyneuen.2016.10.027.

Szabo, S., Tache, Y., Somogyi, A., 2012. The legacy of Hans Selye and the origins of stress research: a retrospective 75 years after his landmark brief "letter" to the editor # of nature. Stress 15 (5), 472–478.

Takai, N., Yamaguchi, M., Aragaki, T., Eto, K., Uchihashi, K., Nishikawa, Y., 2004. Effect of psychological stress on the salivary cortisol and amylase levels in healthy young adults. Arch. Oral Biol. 49 (12), 963–968.

Wang, J., Korczykowski, M., Rao, H., Fan, Y., Pluta, J., Gur, R.C., McEwen, B.S., Detre, J.A., 2007. Gender difference in neural response to psychological stress. Soc. Cogn. Affect. Neurosci. 2 (3), 227–239. http://dx.doi.org/10.1093/scan/nsm018.

Weisman, O., Chetouani, M., Saint-Georges, C., Bourvis, N., Zagoory-Sharon, O., Delaherche, E., Cohen, D., Feldman, R., 2015. Dynamics of non-verbal vocalizations and hormones during father-infant interaction. IEEE Trans. Affect. Comput. http://dx.doi.org/10.1109/TAFFC.2015.2478468.

Wijsman, J., Grundlehner, B., Liu, H., Hermens, H., Penders, J., 2011. Towards mental stress detection using wearable physiological sensors. In: 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2011, pp. 1798–1801.

Wray, N.R., Lee, S.H., Mehta, D., Vinkhuyzen, A.A.E., Dudbridge, F., Middeldorp, C.M., 2014. Research review: polygenic methods and their application to psychiatric traits. J. Child Psychol. Psychiatr. 55, 1068–1087. http://dx.doi.org/10.1111/jcpp.12295.

Zanarini, M., Frankenburg, F.R., 2007. The essential nature of borderline psychopathology. J. Personal. Disord. 21 (5), 518–535. http://dx.doi.org/10.1521/pedi.2007.21.5.518.

Zhang, Z., Girard, J.M., Wu, Y., Zhang, X., Liu, P., Ciftci, U., 2016. Multimodal Spontaneous Emotion Corpus For Human Behavior Analysis. pp. 3438–3446.

Zhou, G., Hansen, J., Kaiser, J.F., 2001. Nonlinear feature based classification of speech under stress. IEEE Transac. Speech and Audio Process. 9 (3), 201–216. http://dx.doi.org/10.1109/89.905995.