

# Design and Validation of a Mental and Social Stress Induction Protocol Towards Load-invariant Physiology-based Stress Detection

Camille Jeunet<sup>1,2</sup>, Fabien Lotte<sup>1</sup> and Christian Mühl<sup>1</sup>

<sup>1</sup>Inria Bordeaux Sud-Ouest / LaBRI, 200 rue de la Vieille Tour, 33405 Talence, France

<sup>2</sup>Laboratoire Handicap & Système Nerveux, Université Bordeaux Segalen, 146 Rue Leo Saignat, 33076 Bordeaux, France

**Keywords:** Mental Stress, Psychosocial Stress, Cognitive Workload, Physiological Computing, ECG, GSR, EEG.

**Abstract:** Stress is a major societal issue with negative impacts on health and economy. Physiological computing offers a continuous, direct, and unobtrusive method for stress level assessment and computer-assisted stress management. However, stress is a complex construct and its physiology can vary depending on its source: cognitive workload or social evaluation. To study the feasibility of physiology-based load-invariant *psychosocial* stress-detection, we designed a stress-induction protocol able to independently vary the relevant types of psychophysiological activity: mental and psychosocial stress. Here, we validate the efficacy of our protocol to induce psychosocial and mental stress. Our participants (N=24) had to perform a cognitive task associated with two workload conditions (low/high mental stress), in two contexts (low/high psychosocial stress), during which we recorded subjects' self-reports, behaviour, physiology and neurophysiology. Questionnaires showed that the subjectively perceived level of stress varied with the psychosocial stress induction, while perceived arousal and mental effort levels vary with mental stress induction. Behaviour and physiology further corroborated the validity of our protocol. Heart rate and skin conductance globally increased after psychosocial stress induction relative to the non-stressful condition. Moreover, we demonstrated that higher workload tasks (mental stress) led to decrease in performance and a marked increase of heart rate.

## 1 INTRODUCTION

Stress is a universal societal issue, affecting both economy and health. Thus, it is easy to understand why many people invest in finding ways to deal with stress (Regehr et al., 2013): how to help people manage their stress is becoming a major preoccupation in many countries. Computer-assisted stress management is one way to support coping with stress. However, it requires reliable stress level assessment (van den Broek and Westerink, 2012).

Besides psychological questionnaires, many devices are available to assess stress levels. They measure stress-related physiological markers such as heart rate, skin conductance or blood pressure, which are increased during a stressful episode (see Section 1.1). The availability of cheap sensor technology and small, portable computing devices allows to automatically and continuously monitor the level of stress in every-day contexts, as during driving (Healey and Picard, 2005) and work (Kusserow et al., 2012), or in clinical contexts (Hogervorst et al., 2013).

However, there are still several challenges that

have to be addressed to be able to successfully monitor stress levels with physiological sensors. One of the most notorious is the definition of the relationship between physiological measurements and the psychophysiological construct of stress. For example, it is known that modifications of the above physiological markers are characteristic of psychosocial stress, but not exclusively affected by it (Dickerson and Kemeny, 2004). For example cardiovascular measures, like heart rate and its variability, are known to respond to stress as well as to high cognitive workload or to exciting situations. The same is true for several frequency bands in the electroencephalogram (EEG): for example the alpha frequency band has been shown to covary with stress/relaxation, but is also known to respond strongly to sensory stimulation, attention and cognitive workload (see Section 1.1). Therefore, the relationship between measurements and psychophysiological constructs is a complex many-to-many mapping for which great care has to be taken in the aim of finding the right mapping parameters to avoid confusion between episodes of psychosocial stress and those of high cognitive workload, i.e., of mental stress.

To enable the creation of systems that are capable of differentiating activation of body and brain due to psychosocial stress from those due to mental stress (associated to high cognitive workload), it is necessary to ensure psychophysiological validity of the inference, as well as the diagnosticity and generality of the detection system (Fairclough, 2009). To tackle this essential requirement of physiological computing, we devised an affect induction protocol that manipulates mental and bodily activation due to psychosocial and mental stress independently. The data gathered with it deliver the ground truth necessary to calibrate reliable stress classifiers. Here, we present the protocol and subjective (self-assessments) and objective (physiological and behavioural measures) evidence of its validity.

This paper is organized as follows: concepts of stress and related work on physiological computing as well as our hypotheses are presented in the remainder of the introduction, Section 2 presents material and methods, then results are presented in Section 3 and discussed in Section 4; finally, we conclude our work with a perspective on future studies in Section 5.

## 1.1 Stress Responses

The psychophysiological concept of 'stress' was introduced in 1936 by (Selye, 1936) to describe 'the non-specific response of the body to any demand for change'. In that sense, it is an organism's response to an environmental situation or stimulus perceived negatively - called a 'stressor' - which can be real or imagined, that taxes the capacities of the subject, and thus has an impact on the body's homeostasis (that is to say that the constants of the internal environment are modified). To face the demand (i.e., to restore homeostasis), two brain circuitries can be activated during a 'stress response cascade' (Sinha et al., 2003; Dickerson and Kemeny, 2004; Taniguchi et al., 2009): the sympatho-adrenomedullary axis (SAMA, also called the noradrenergic circuitry) and the hypothalamus-pituitary gland-adrenal cortex axis (HPAa). On the one hand, the SAMA induces the release of noradrenaline which allows immediate physical reactions (such as increased heart rate and skin conductance, or auditory and visual exclusion phenomena) associated with a preparation for violent muscular action (Dickerson and Kemeny, 2004). On the other hand, HPAa activation (which is lower) results in the releasing of cortisol, the purpose of which is to redistribute energy in order to face the threat. Thus, more energy is allocated to the organs that need it most (brain and heart), while non-necessary organs for immediate survival (reproductive, immune and di-

gestive systems) are inhibited. This stress response cascade ends when homeostasis is restored. However, stress can be of different types, such as physical, psychological or psychosocial (Dickerson and Kemeny, 2004), each kind of stress being associated with a specific response (contrary to what has previously been said in Selye's definition of stress). Indeed, physical stress, induced by extreme temperatures or physical pain for example, is associated with an increase of heart rate (Loggia et al., 2011), skin conductance (Boucsein, 1992; Buchanan et al., 2006) and subjective stress ratings but with only a weak cortisol response (Dickerson and Kemeny, 2004). These results suggest that this kind of stress induces an activation of the SAMA but only a weak activation of the HPAa. Mental (or psychological) stress, associated with difficult cognitive tasks (i.e., high workload), uncontrollability or negative emotions triggers a weak release of cortisol (weak HPAa activation), but strong effects on heart rate and skin conductance (strong SAMA activation) (Boucsein, 1992; Reinhardt et al., 2012). Finally, psychosocial stress, triggered by a social evaluation threat (that is to say a situation in which the person's own estimated social value is likely to be degraded) added to uncontrollability (in particular during the Trier Social Stress Task (TSST) (Kirschbaum et al., 1993)) has been shown to induce a strong activation of both the SAMA (Hellhammer and Schubert, 2012) and the HPAa (Dickerson and Kemeny, 2004).

Thanks to research in the field of neurosciences, and consistently with the notion of neural response systems, we know that stress also has strong electroencephalographic correlates. One of the most prominent correlates of psychosocial stress is found in the alpha band, and specifically in brain asymmetry. Tops et al. (Tops et al., 2006) proposed that cortisol administration (which simulates a stressful situation) leads to a global decrease of cortical activity (except for the left frontal cortex in which activity is increased). However, other studies (Lewis et al., 2007; Hewig et al., 2008) showed that stress was associated with a higher activity in the right hemisphere, and that right hemisphere activation was correlated with negative affect. For Crost and colleagues (Crost et al., 2008), the explanation of these conflicting results would be that an association between EEG-asymmetry and personality characteristics may only be observed in relevant situations to the personality dimensions of interest. As for the other physiological responses, different types of stress share similar neurophysiological responses. For mental stress for example, we know that the alpha band plays a role as well (Brouwer et al., 2012). Given the overlap in the response characteristics between different types of

stress, it is essential to explore these types of stress in context, that is to say in one experiment.

To summarise, stress is a process composed of five phases: -1- an environmental demand that disrupts homeostasis, -2- a personal perception of this demand by the person, -3- a stressor-specific activation of the stress-response, -4- behavioural consequences of this activation, -5- a return to homeostasis once the threat is overcome, or exhaustion if the body has not enough resources to face the demand. Depending on the type of stressor, different types of stress occur, sharing mechanisms and overlapping in response characteristics. To identify valid psychophysiological inferences in every-day stress detection systems, these types of stress have to be experimentally delineated.

## 1.2 Physiological Computing and Stress

Several works in the domain of physiological computing have addressed the measurement of various types of stress from physiological and neurophysiological sensors. For example, Healey has shown that during different driving phases, various levels of stress can be automatically inferred from physiological sensors, assessing cardiovascular and respiratory responses (Healey and Picard, 2005). Hogervorst and colleagues showed that in the context of an impending eye surgery, physiological signals such as heart rate can be used to classify stress (Hogervorst et al., 2013). Khosrowabadi and colleagues showed that stress emerging during an examination period can be identified from the EEG (Khosrowabadi and Quek, 2011). Riera found that neurophysiological signals can differ between the different levels of stress elicited in an interview setting (Riera et al., 2012).

As these studies show, the stress response can be automatically detected using information recorded from the somatic, autonomous, and central nervous systems, which therefore pose a potential basis for stress detection systems. However, as laid out in the former section, physiological stress responses due to psychosocial stress and those due to mental demand resemble each other. This is especially relevant for the construction of stress detection systems, since in every-day life and work mental demand and psychosocial stress are often co-occurring: mental task demand might lead to psychosocial stress due to the associated evaluation and psychosocial stress often is related to certain (demanding) tasks that have to be conducted. Therefore, care has to be taken to ensure the validity of the psychophysiological inference offered by a physiology-based stress detection method. As stressed by Fairclough (Fairclough, 2009), psychophysiological validity can be ensured by a number

of approaches, including standardised experimental stimuli and tasks for the induction of the psychophysiological constructs in question, and the assessment of subjective and objective measures, such as observable behaviour or physiological signals known to respond to the psychophysiological construct in question. Returning to the above mentioned studies, no distinction between the type of stress involved has been made. However, such distinction is critical for systems that are supposed to be sensitive to one specific type of stress, while being invariant for others. That is why it is important to devise a protocol that allows to vary psychosocial and mental stress independently.

Our long-term goal is the development of a system that can be used during private or professional work, which is able to differentiate psychophysiological activation due to stress compared to that due to cognitive workload (mental stress). Therefore, we devised a stress induction protocol that would take the following requirements into account: 1) enable recording during actual tasks, not during resting periods (ecological validity), 2) manipulate the context of the task, but maintain same sensory stimulation and task during measurement (avoiding confounds), 3) introduce different levels of cognitive workload during stressful and non-stressful conditions (enabling the creation of a classifier specific to psychosocial stress and invariant to mental stress). In line with Fairclough's requirements, we combine two recognised experimental paradigms to induce psychosocial and mental stress and validate this protocol via subjective ratings and objective measures from behaviour and physiology.

## 1.3 Hypotheses

To ensure the validity and efficiency of the affective manipulation (i.e., psychosocial stress) within our stress induction protocol, we analyzed participants' self-assessments (State-Trait Anxiety Inventory, arousal scale of the Self-Assessment Maneken) to see if subjective experience of anxiety and arousal was increasing in the stressful condition of our protocol relative to the control condition. For a more objective evaluation, we also analysed the physiological (heart rate and skin conductance) differences according to the condition (stress vs. no-stress). According to the literature provided in Section 1.1, we were expecting an increase of the heart rate and skin conductance during the stressful condition. To ensure the validity and efficacy of the cognitive workload manipulation (i.e., mental stress), we analysed participants' self-assessments (Rating Scale Mental Effort, arousal scale of the Self-Assessment Maneken) expecting to find an increase in perceived cognitive load

and arousal during high cognitive workload compared to low cognitive workload. Furthermore, we expected a decrease of task performance with higher workload. Finally, we also hypothesised an increase of heart rate with high workload (see Section 1.1).

## 2 METHODS

### 2.1 Participants

12 female and 12 male participants were recruited for our experiment. The participants were between 18 and 54 years old, with a mean age of 24.7 (SD: 7.9), and all except 4 were right-handed. Educations varied between high school degree and Ph.D., with a mean education of high school + 3.1 years (SD: 2.4). To be admitted, people had to be at least 18 and had to sign an informed consent. Furthermore, we had criteria of non inclusion: bad vision, heart, neurological and psychological diseases, and affective troubles. Moreover, people were asked to schedule on a day and at a time at which they expected to feel good, not too tired. Finally, we asked them not to drink coffee and tea less than two hours before the experiment.

### 2.2 Material

For our recordings, we used the following sensors (see Figure 1): electroencephalogram (EEG, 28+8 -in a 10/20 system without T7, T8, Fp1 and Fp2- active electrodes -g.LADYbird- with g.USBamp), electrocardiogram (ECG, 2 active electrodes), electromyogram (EMG, 2 active electrodes), electrooculogram (EOG, 4 active electrodes), breath belt (SleepSense), pulse-sensor (g.PULSEsensor), and a galvanic skin response sensor (g.GSRsensor). The workload task was designed in the Presentation software (Neurobehavioral Systems, [www.neurobs.com/presentation](http://www.neurobs.com/presentation)) and EEG signals were recorded and visually inspected with Open ViBE (Renard et al., 2010).

Subjects were first asked to sign an informed consent and to fill in three questionnaires. The first one concerned personal characteristics (such as gender, age and education). The second and the third ones were State-Trait Anxiety Inventory (STAI) form Y-A and Y-B (Spielberger et al., 1970) (see 2.2.1 for more details). Then, all the sensors were put into place and a three minute baseline was recorded. From this point, as we wanted to counterbalance the conditions (not to have any order effect), we set up four scenarios (see Figure 2) composed of two blocks each, separated by a STAI form Y-A questionnaire. Therefore, we randomly began with either relaxation or stress

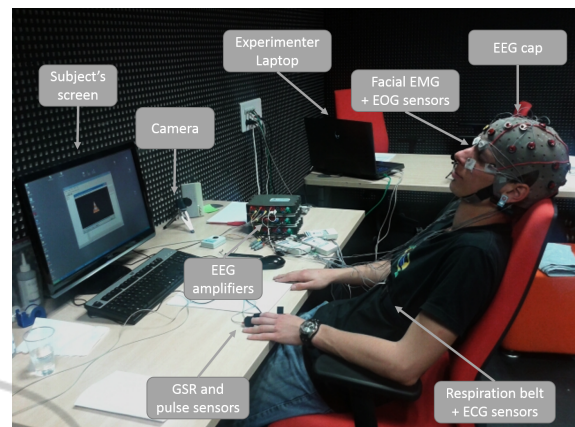


Figure 1: Illustration of our set up.

induction, and we randomly started with either low workload (0) or high workload (2). During both the first and the second block, the subject did each workload condition six times (low/high) ( $6 \times 2 \times 2 = 24$  min per block), with a short break after six tasks. Once the two blocks were completed, the subject filled in the STAI form Y-A one last time before the sensors were taken off. Finally, the participant was debriefed about the aim of the experiment.

#### 2.2.1 Neuropsychological Stress Evaluation

In order to measure the level of anxiety of the subjects, the "State Trait Anxiety Inventory" (Spielberger et al., 1970) was used. This test is composed of two scales of 20 propositions each: STAI form Y-A and STAI form Y-B. STAI form Y-A score increases in a stressful situation. It is a good indicator for transitory modifications of the level of stress. STAI form Y-B evaluates the clinical anxiety (trait), and thus allows to recognize generally anxious people (who have higher scores).

#### 2.2.2 Psychosocial Stress and Relaxation Inductions

In order to induce stress, a 15 minutes long stress-induction protocol based on the Trier Social Stress Task (or TSST, (Kirschbaum et al., 1993)) which is a validated protocol, was set up. This stress induction protocol is composed of three parts and requires the participation of three people, "the committee", who are presented as being body language experts. In the first part, a member of the committee tells the subject he has five minutes to prepare a fake job interview for a position as a teacher. During the second part, the committee asks the person to do this job interview and to speak about himself for 5 minutes. They tell the subject that he is filmed (for a future behavioural



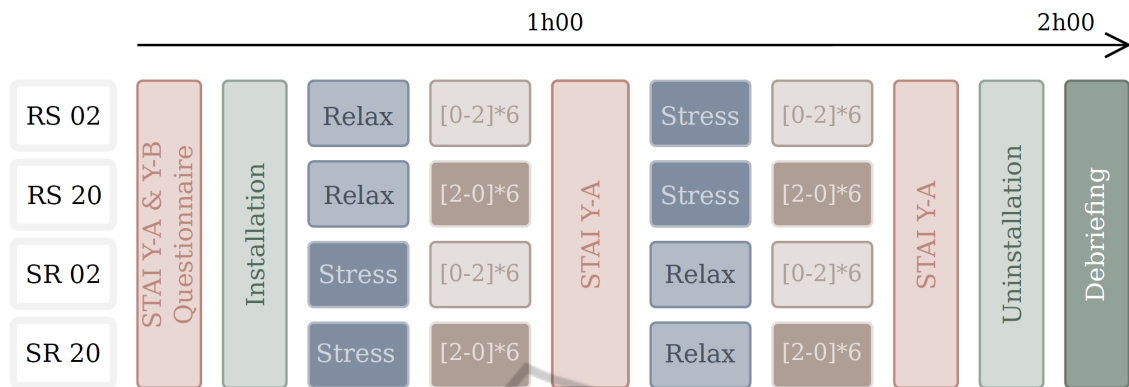


Figure 2: Chronology of our Experimental Protocol: each row represents a different scenario (R=Relax; S=Stress; 0=0-back; 2=2-back).

analysis) and they take notes. The committee acts serious and neutral/unresponsive towards the subject. The third part is an arithmetic task lasting 3 minutes. The subject has to count backwards from 2083 to 0 by steps of 13 and to begin again at any mistake or hesitation. At the end of this protocol, in order to keep the stress level high, the comity tells the subject he will be filmed during the workload tasks and that he will have to do another interview, which will be longer, and a self-evaluation after it. Furthermore, during the experiment, participants have to perform cognitive tasks, followed by feedback corresponding to their performance. During the stress condition, the score shown to the participant was automatically decreased by 5 to 10%. Thus, this protocol includes psychosocial stress and uncontrollability in order to maximise the chances of triggering a stress response in all participants (Dickerson and Kemeny, 2004). On the other hand, the goal of the relaxation induction was to create a condition in which participants would be able to relax. In this last part, they were allowed to either rest in silence, or select relaxing music or videos (Krout, 2007).

### 2.2.3 Workload Tasks

We used the n-back task (Kirchner, 1958) to manipulate mental stress, as it was easy to modify the workload without changing any other parameter (such as visual stimulation or motor behaviour). Indeed, 0-back (low workload) and 2-back (high workload) are very similar. In both, sixty white letters (font-size 48) appear (during 500ms) one after the other (with a 1500ms break between them) on a black background. Thus, each task lasts two minutes. Among these letters, 25% are targets. In both tasks, when the letter that appears is a target, the subject has to do a left click, and a right click otherwise (like this, there will not be a big difference at the motor level between targets and non target letters: this motor dif-

ference could have induced unexpected differences in the EEG). Thus, for the 0-back task, that is to say the low workload condition, the target is the letter "X": each time an "X" appears, the subject has to do a left click, and in all the other cases he has to do a right click. For the 2-back task, the high workload condition, the subject has to do a left click when the letter that appears is the same as two letters before (for example, if the sequence "C A C" appeared, the second "C" would be a target). At the end of each task, the subject has to report his perceived level of arousal (on a scale from 1 to 9 of the Self-Assessment Maneken, SAM) and the spent mental effort (Rate Scale of Mental Effort, RSME) (Brouwer et al., 2012). Finally, a screen with the subject's performance at the prior task (number of targets correctly identified) appears.

## 2.3 Variables and Factors

In this study, we analysed the effects of two factors (with two possible values each) on different dependent variables (DV). The first factor is the psychosocial stress level: condition "stress" (after the stress induction) vs. condition "no-stress" (after relaxation induction); the second factor is the mental stress level: low workload (low-WL, 0-back task) vs. high workload (high-WL, 2-back task). Thus, we had four experimental conditions: stress/low-WL, stress/high-WL, no-stress/low-WL and no-stress/high-WL. The DVs are: STAI form Y-A score, subjective arousal evaluation, subjective task difficulty evaluation, tasks performance, heart rate and skin conductance.

## 2.4 Data Analyses

### 2.4.1 Self-assessment Data

To investigate the effect of psychosocial stress induction on the state-trait anxiety inventory, we computed

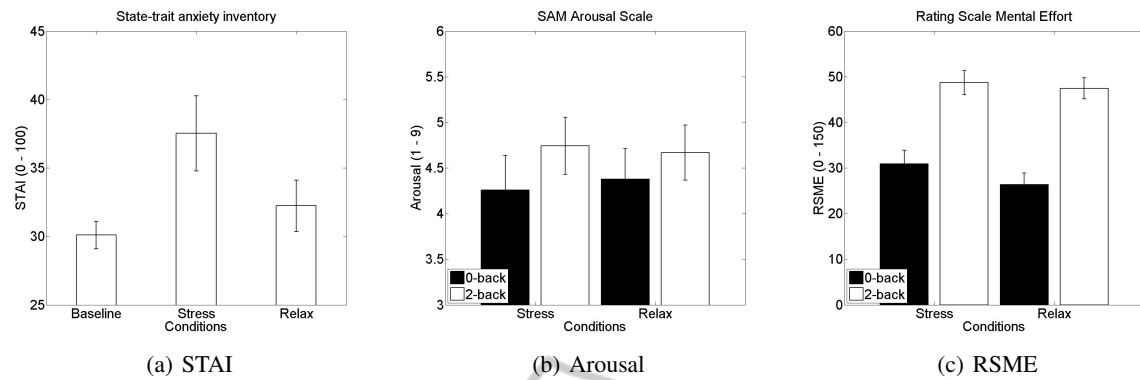


Figure 3: Mean and standard error of mean of subjective stress level assessments : (a) STAI form Y-A scores, (b) SAM arousal scale and (c) RSME. (a) shows significant increase of perceived stress during the stressful condition compared to the baseline and the relax condition. (b) and (c) show an increase of perceived arousal and mental effort for the 2-back compared to 0-back task.

an ANOVA with the self-ratings in the 3 factor-levels (baseline, after relaxation, after stress induction). To test the effect of psychosocial and mental stress manipulation, we conducted 2 (stress)  $\times$  2 (workload) ANOVAs for the averaged-over-blocks ratings (one value per subject for stress/low workload, stress/high workload, relaxation/low workload, relaxation/high workload blocks) on the arousal scale of the SAM and on the RSME.

#### 2.4.2 Behavioural Data

To investigate the effects of the experimental manipulations on behaviour, we calculated the performance per block based on the number of true positive ( $TP$ ), true negative ( $TN$ ), false negative ( $FN$ ), and false positive ( $FP$ ) responses resulting from the button presses within the n-back task (left click for targets, right click for non-targets) using the following equation:  $Perf = (TP + TN) \div (TP + TN + FP + FN)$ . As for ratings, we analysed the data then in a 2 (stress)  $\times$  2 (workload) ANOVA.

#### 2.4.3 Physiological Data

Physiological responses were analysed with respect to heart rate (HR) and galvanic skin response (GSR). Before applying statistical methods, the GSR data were pre-processed by extracting the GSR value ( $\mu S$ ) for each block and then averaging these values over the blocks described above. The ECG signal was band-pass filtered between 5 and 200 Hz, applying a notch-filter 48-52 Hz to reduce power line noise, before HR for each of the blocks was extracted using the BioSig toolbox<sup>1</sup>. As for the former analyses, we then analysed the data in a 2 (stress)  $\times$  2 (work-

load) ANOVA. We are reporting data as significant if  $p < 0.05$  and as trend if  $p < 0.1$ .

## 3 RESULTS

### 3.1 Subjective Indicators

#### 3.1.1 State-trait Anxiety Inventory

Each subject filled in three "STAI form Y-A" (state) questionnaires: at the beginning ( $STAI_{BL}$ ) of the experiment, and after each of the stress conditions (stress:  $STAI_S$ ; no-stress:  $STAI_R$ ) (see Figure 3(a)). Three datasets were excluded due to incompleteness. An ANOVA ( $N = 21$ ) with 'baseline', 'stress', and 'relax' factor levels showed a significant difference of perceived anxiety between the conditions ( $F(2,20) = 3.6225$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.108$ ). We conducted a post-hoc analysis using directed paired t-tests according to the hypothesis that subjectively perceived stress increases due to the stress induction procedure relative to baseline and relaxation condition. The results show the expected effect (see Figure 3(a)).  $STAI_S$  ( $M_S = 37.5$ ,  $SD_S = 12.6$ ) scores are significantly higher ( $t(20) = 2.87$ ,  $p = 0.005$ ) than  $STAI_{BL}$  scores ( $M_{BL} = 30.1$ ,  $SD_{BL} = 4.6$ ) and also significantly higher ( $t(20) = 2.37$ ,  $p = 0.014$ ) than  $STAI_R$  scores ( $M_R = 32.2$ ,  $SD_{BL} = 8.6$ ). We found no difference between  $STAI_R$  and  $STAI_{BL}$  ( $t(20) = 1.27$ ,  $p = 0.11$ ).

#### 3.1.2 Perceived Arousal and Mental Effort

We asked the subjects after each block to rate their arousal on the respective scale of the Self-Assessment Maneken (see Figure 3(b)) and to rate the invested mental effort on the Rating Scale Mental Effort (see

<sup>1</sup><http://biosig.sourceforge.net/>

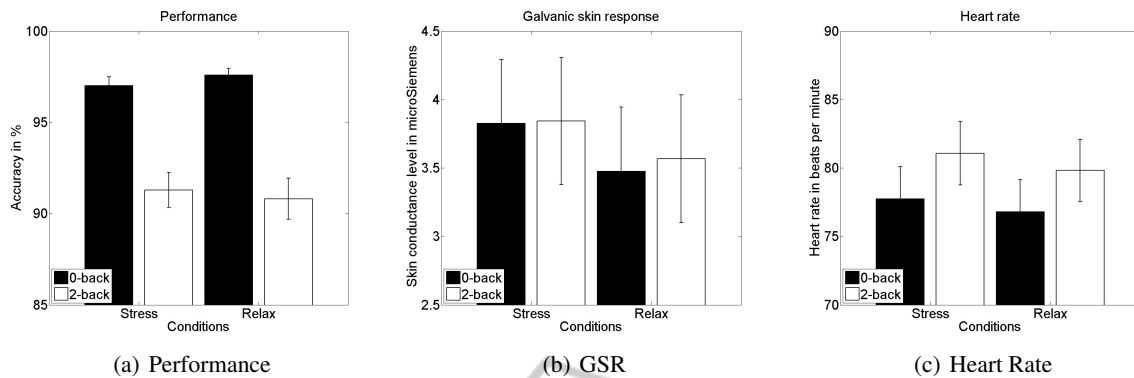


Figure 4: Mean and standard error of mean of objective stress level assessments : (a) Performance, (b) Galvanic Skin Response and (c) Heart Rate. (a) depicts a lower performance in 2-back compared to 0-back tasks. (b) shows that GSR significantly increased only for the "stress" condition compared to the "relax" condition. (c) depicts a significant increase of heart rate during high workload compared to low-workload tasks.

Figure 3(c)). Two datasets were excluded due to incompleteness. We submitted the data of each scale to a 2 (stress)  $\times$  2 (workload) ANOVA. Regarding subjectively perceived arousal, we only found a main effect of the workload manipulation ( $F(1,21) = 4.444$ ,  $p = 0.047$ ,  $\eta_p^2 = 0.175$ ) with higher perceived arousal for the 2-back task ( $M_{2-back} = 4.7$ ,  $SD_{2-back} = 1.4$ ) compared to the 0-back task ( $M_{0-back} = 4.3$ ,  $SD_{0-back} = 1.7$ ). Regarding the subjectively perceived workload, we only found a main effect of the workload manipulation ( $F(1,21) = 63.216$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.751$ ) with higher perceived effort for the 2-back task ( $M_{2-back} = 48.1$ ,  $SD_{2-back} = 11.5$ ) compared to the 0-back task ( $M_{0-back} = 28.6$ ,  $SD_{0-back} = 12.9$ ). Although the interaction between workload and stress did not reach significance ( $F(1,21) = 2.345$ ,  $p = 0.141$ ), an exploratory post-hoc paired t-test ( $t(21) = 2.7$ ,  $p = 0.013$ ) showed greater perceived mental effort during low workload when exposed to psychosocial stress than during 'relaxation' condition.

## 3.2 Objective Indicators

### 3.2.1 Performance

For the analysis of task performance the accuracy for each task block (see Figure 4(a)) was submitted to a 2 (stress)  $\times$  2 (workload) ANOVA. Two datasets were excluded due to incompleteness. As for the subjective indicators of perceived arousal and effort, we found a main effect of workload manipulation ( $F(1,21) = 65.251$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.757$ ) with higher accuracy for the simple 0-back task ( $M_{0-back} = 97.3$ ,  $SD_{0-back} = 2.0$ ) compared to the hard 2-back task ( $M_{2-back} = 91.1$ ,  $SD_{2-back} = 4.8$ ).

### 3.2.2 Physiological Sensors

We submitted mean skin conductance level and heart rate for each block to a 2 (stress)  $\times$  2 (workload) ANOVA. Four datasets were excluded due to incompleteness. For heart rate analysis a further dataset was excluded due to its low signal quality. As for GSR (see Figure 4(b)), we found an increase of skin conductance ( $F(1,19) = 4.4806$ ,  $p = 0.048$ ,  $\eta_p^2 = 0.191$ ), indicating higher sympathetic arousal during the "stress condition" ( $M_S = 3.83$ ,  $SD_S = 2.05$ ) compared to the "relax" condition ( $M_S = 3.52$ ,  $SD_S = 2.07$ ). Skin conductance was also observed to increase (although not significantly) during the high workload condition compared to the low workload condition. For HR (see Figure 4(c)), we found a trend towards an increase of heart rate ( $F(1,18) = 3.2123$ ,  $p = 0.089$ ,  $\eta_p^2 = 0.151$ ), indicating higher sympathetic arousal during the stress condition ( $M_S = 79.41$ ,  $SD_S = 10.23$ ) compared to the relax condition ( $M_R = 78.30$ ,  $SD_R = 10.08$ ). Furthermore, we found a highly significant effect of workload manipulation on HR ( $F(1,18) = 36.1431$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.667$ ), with a higher heart rate for the more challenging 2-back task ( $M_{2-back} = 80.4$ ,  $SD_{2-back} = 9.89$ ) compared with the relatively easy 0-back task ( $M_{0-back} = 77.27$ ,  $SD_{0-back} = 10.19$ ).

## 4 DISCUSSION

We set out to devise an experimental protocol allowing to manipulate psychosocial and mental stress independently and presented evidence for its validity. The score of the STAI (anxiety state inventory) showed a significant response to psychosocial stress manipulation, as STAI scores were significantly

higher in the 'stress' condition compared to 'baseline' and 'relaxation' conditions. Given that the form was filled in at the end of the 'stress' and 'relaxation' conditions, after the workload tasks, the still increased scores in the 'stress' condition are evidence for a high psychosocial stress level throughout the whole 'stress' condition. Furthermore, regarding the arousal scale of the SAM, we found no evidence for a higher perceived arousal during the 'stress' condition. The lacking effect of the stress manipulation on that scale might be explained by a domination of the workload manipulation on arousal, as it was given with the rating scale on mental effort after each task. Furthermore, a post-hoc analysis showed that in the low-workload condition (i.e., 0-back task), perceived mental effort (on the RSME scale) was higher in the 'stress' condition than in the 'relaxation' condition. This result suggests that perceived mental effort increases with psychosocial stress level, and thus indicates that more cognitive resources are needed to perform a task in a highly stressful situation. It is in line with the dynamic adaptability theory of stress (Szalma and Teo, 2012) which states that psychosocial stress induces an increase in perceived workload. The fact that this pattern is not observable in the high workload condition (i.e., 2-back task) may be due to a ceiling effect: the fact that the task itself induces a high level of perceived mental effort decreases the probability of increased mental effort due to psychosocial stress. This ceiling effect could be responsible of the apparent (but not significant) interaction in Figure 3(c). Regarding the mental stress manipulation, the arousal scale showed increased arousal for the high workload condition, mirroring the higher perceived mental effort for this condition on the RSME scale. Moreover, the performance decreased for the 2-back task. These results deliver further evidence for the validity of the mental stress manipulation, since they show that the 2-back task is much more difficult (requires a significantly higher workload level) than the 0-back task.

Furthermore, the physiological signals delivered evidence for the validity of the psychosocial as well as of the mental stress manipulation. The elevated skin conductance level and a trend towards a higher heart rate during the stressful task-context corroborated the effect found for the subjective stress ratings. Finding elevated sympathetic nervous system activity after the stress induction phase (TSST) and throughout the task period is strong evidence that we were able to stretch the phase of increased psychosocial anxiety over the interview itself, using instructions and camera setup (Hellhammer and Schubert, 2012). Furthermore, a higher heart rate during difficult (2-back) compared

to simple (0-back) task conditions supported the validity of the mental stress induction via task difficulty. Finally, we found no significant interaction effects, neither for subjective (STAI, arousal, RSME) nor for objective indicators (heart rate, GSR). This is convenient, as it practically excludes the possibility of differences between conditions others than the intended ones (i.e., psychosocial and mental stress) and thus may indicate that the goal has been achieved: we designed and validated a stress induction protocol allowing to vary mental and psychosocial stress independently.

## 5 CONCLUSIONS

The presented results show that we managed to design and validate a stress induction protocol which allows the experimenter to vary mental stress and psychosocial stress independently. However, although we devised a strict protocol to avoid confusion with other behavioural parameters as well as order effects, this protocol remains very simple and not very ecological. That is why future work will focus on continuous classification (instead of only two classes: stress/relax and easy/difficult) which would be more relevant for real-life applications. Finally, two other goals will have to be achieved : 1) EEG data analyses will be performed in order to study the feasibility of an EEG-based (mental and psychosocial) stress classification; 2) the final goal of this project is to design and validate a load-invariant physiology-based stress detection protocol, which will be used to help people manage their stress at home and at work.

## REFERENCES

- Boucsein, W. (1992). Electrodermal Indices of Emotion and Stress. In *Electrodermal Activity*, chapter 3, pages 369–391. Springer New York.
- Brouwer, A. M., Hogervorst, M. A., van Erp, J. B. F., Heffelaar, T., Zimmerman, P. H., and Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. *International Journal of Neural Engineering*, 9.
- Buchanan, T. W., Etze, J. A., Adolphs, R., and Tranel, D. (2006). The influence of autonomic arousal and semantic relatedness on memory for emotional words. *International Journal of Psychophysiology*, 61:23–26.
- Crost, N. W., Pauls, C. A., and Wacker, J. (2008). Defensiveness and anxiety predict frontal EEG asymmetry only in specific situational contexts. *Biological Psychology*, 78:43–52.
- Dickerson, S. S. and Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration



- and synthesis of laboratory research. *Psychological Bulletin*, 130(3):355–391.
- Fairclough, S. H. (2009). Fundamentals of physiological computing. *Interacting with Computers*, 21:133–145.
- Healey, J. and Picard, R. (2005). Detecting Stress During Real-World Driving Tasks Using Physiological Sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6(2):156–166.
- Hellhammer, J. and Schubert, M. (2012). The physiological response to Trier Social Stress Test relates to subjective measures of stress during but not before or after the test. *Psychoneuroendocrinology*, 37(1):119–24.
- Hewig, J., Schlotz, W., Gerhard, F., Breitenstein, C., Luerken, A., and Naumann, E. (2008). Associations of the cortisol awakening response (CAR) with cortical activation asymmetry during the course of an exam stress period. *Psychoneuroendocrinology*, 33:83–89.
- Hogervorst, M. A., Brouwer, A.-M., and Vos, W. K. (2013). Physiological Correlates of Stress in Individuals About to Undergo Eye Laser Surgery. In *5th IEEE International Conference on Affective Computing and Intelligent Interaction*, Geneva, Switzerland. IEEE Computer Society Press.
- Khosrowabadi, R. and Quek, C. (2011). A Brain-Computer Interface for classifying EEG correlates of chronic mental stress. In *Proceedings of International Joint Conference on Neural Networks 2011 (IJCNN)*, volume 138632, pages 757–762.
- Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, 55(4):352–358.
- Kirschbaum, C., Pirke, K.-M., and Hellhammer, D. H. (1993). The "Trier Social Stress Test": a Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. *Neuropsychobiology*, 28:76–81.
- Krout, R. E. (2007). Music listening to facilitate relaxation and promote wellness: Integrated aspects of our neurophysiological response to music. *The Arts in Psychotherapy*, 34:134–141.
- Kusserow, M., Amft, O., and Troster, G. (2012). Monitoring Stress-arousal in the Wild - Measuring the Intangible. *IEEE Pervasive Computing*, 12(2):28 – 37.
- Lewis, R. S., Weekes, N. Y., and Wang, T. H. (2007). The effect of a naturalistic stressor on frontal EEG asymmetry, stress, and health. *Biological Psychology*, 75:224–239.
- Loggia, M. L., Juneau, M., and Bushnell, M. C. (2011). Autonomic responses to heat pain: Heart rate, skin conductance, and their relation to verbal ratings and stimulus intensity. *Pain*, 152(3):592–8.
- Regehr, C., Glandy, D., and Pitts, A. (2013). Interventions to reduce stress in university students: A review and meta-analysis. *Journal of Affective Disorders*, 148:1–11.
- Reinhardt, T., Schmahl, C., Wüst, S., and Bohus, M. (2012). Salivary cortisol, heart rate, electrodermal activity and subjective stress responses to the Mannheim Multi-component Stress Test (MMST). *Psychiatry research*, 198(1):106–11.
- Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O., and Lécuyer, A. (2010). OpenViBE: An Open-Source Software Platform to Design, Test, and Use Brain-Computer Interfaces in Real and Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 19(1):35–53.
- Riera, A., Soria-Frisch, A., Albajes-Eizagirre, A., Cippresso, P., Grau, C., Dunne, S., and Ruffini, G. (2012). Electro-physiological data fusion for stress detection. *Studies in health technology and informatics*, 181:228–32.
- Selye, H. (1936). A syndrome produced by diverse noxious agents. *Nature*, 138:32.
- Sinha, R., Talih, M., Malison, R., Cooney, N., Anderson, G.-M., and Kreek, M.-J. (2003). Hypothalamic-pituitary-adrenal axis and sympatho-adrenomedullary responses during stress-induced and drug cue-induced cocaine craving states. *Psychopharmacology*, 170:62–72.
- Spielberger, C. D., Gorsuch, R. L., and Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press, Palo Alto.
- Szalma, J. L. and Teo, G. W. L. (2012). Spatial and temporal task characteristics as stress: A test of the dynamic adaptability theory of stress, workload, and performance. *Acta Psychologica*, 139:448–471.
- Taniguchi, K., Nishikawa, A., Sugino, T., Aoyagi, S., Sekimoto, M., Takiguchi, S., Okada, K., Monden, M., and Miyazaki, F. (2009). Method for Objectively Evaluating Psychological Stress Resulting when Humans Interact with Robots. In Kulyukin, V. A., editor, *Advances in Human-Robot Interaction*. InTech China.
- Tops, M., van Peer, J. M., Wester, A. E., Wijers, A. A., and Korf, J. (2006). State-dependent regulation of cortical activity by cortisol: An EEG study. *Neuroscience Letters*, 404:34–39.
- van den Broek, E. L. and Westerink, J. H. D. M. (2012). Biofeedback Systems for Stress Reduction - Towards a Bright Future for a Revitalized Field. In *International Conference on Health Informatics (HEALTH-INF 2012)*, pages 499–504.