Inducing and Assessing Acute Mental Stress in Controlled Conditions: Topical Review and Guidelines for Effective Experimental Protocols

IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 14 December 2024, accepted 8 January 2025, date of publication 13 January 2025, date of current version 16 January 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3528518

TOPICAL REVIEW

Inducing and Assessing Acute Mental Stress in Controlled Conditions: Topical Review and Guidelines for Effective Experimental Protocols

MATTEO RAGGI[1], **LEE J. MOORE**[2], **AND LUCA MESIN**[1]
[1]Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Turin, Italy
[2]Department for Health, University of Bath, Bath, BA2 7AY Somerset, U.K.

Corresponding author: Matteo Raggi (matteo.raggi@polito.it)

This work was supported in part by the European Union–NextGenerationEU.

ABSTRACT Mental health is influenced by the fast-paced nature of life. In this scenario, stressful events play an important role. Extensive research has been carried out to develop non-invasive devices for stress detection, which primarily use physiological data and, more recently, artificial intelligence algorithms. When developing either a new device or algorithm, tests in controlled environments are preferred, because of the supervision of possible confounding factors while running the experiments. However, because of the extremely subjective perception of stress, the characteristics of the investigated samples, and the conditions under which the experiment is conducted, the data may not be representative of a perceived stressful condition, leading to biases. Given the importance of reliable experimental protocols for stress induction, especially if cortisol level is not monitored, this work aims to present approaches for inducing and assessing acute mental stress in controlled conditions, analyzing the problem from engineering and psychological perspectives. All the phases of the experimental protocol are discussed, examining both the factors that could induce stress and the assessment tools, like questionnaires and physiological signals. The analysis of the latter will be focused on the exogenous factors that may compromise the measures, providing solutions for their mitigation. With this work, researchers with different backgrounds can improve the efficacy of their studies, limiting biases and misleading results.

INDEX TERMS Stress monitoring, stress assessment, state anxiety, experimental protocols.

I. INTRODUCTION

The hectic pace of life and the pressures we are subject to daily affect our mental health, which is a part of the ''health'' according to World Health Organization [1]. In this scenario, a significant role is played by psychological stress, which manifests when the demands of a situation are perceived to overwhelm the resources we have available to cope [2]. Notably, we perceive stress when facing high-demanding situations, such as financial instabilities and divorces. Unfortunately, this is increasingly true nowadays. A survey by Mental Health UK published in 2024 pointed out that more than 90 % of the participants experienced high or extreme pressure or stress over the previous year [3].

The stress response can be either acute or chronic, depending primarily on the duration of the stressful events [4]. With particular attention to the chronic condition, it has been demonstrated that long exposure to stress may compromise health in older or unhealthy individuals [5]. The onset of cardiovascular diseases [6], metabolic disorders [7] and anxiety [8] are just a few examples of the stress-induced

The associate editor coordinating the review of this manuscript and approving it for publication was Muammar Muhammad Kabir.

diseases, with a direct cost for national health systems and companies, the latter estimated around £28 billion only in the United Kingdom [9].

Given the significant impact of stress and its negative effects, extensive research has been carried out in the last two decades in this regard. A tangible result is the worldwide diffusion of wearable devices [4], [11], [12], which acquire both physical and physiological signals, useful for stress detection, in an easy, reliable and inexpensive way. The heart rate and its variability (HRV) [13], [14], [15] and the electrodermal activity (EDA) [16], [17], [18] are just a few examples of the signals that can be collected by either smartwatches or wristbands, commonly available on the market. The Fitbit Sense 2 smartwatch, along with Whoop 4.0 and the Empatica E4 wristbands, are just a few examples.

In this context, machine learning (ML) [11], [19], [20] and deep learning (DL) algorithms [12], [21], [22], [23] are playing a pivotal role, leading to stress estimation with high accuracy with either unimodal or multimodal approaches [11], [19], [24]. However, despite the great effort from ongoing research, the analysis of cortisol (CORT) levels from biological samples is still today the gold standard for this purpose [25], albeit the procedure is expensive and not immediate [26], limiting its use in real-time applications [27], [28].

Even though the aim is to appraise the users' stress levels in everyday life (i.e., outside the laboratory), testing a new solution, either hardware or software, in controlled conditions is always a good practice; in fact, the possible confounding factors could be limited and the conditions where the experiments take place can be monitored. In this scenario, stressful events need to be artificially recreated. Different approaches could be considered, but cognitive, audiovisual or speech tasks are certainly the most adopted. In this regard, a consistent number of public datasets for ML and DL algorithm development exploit some of the mentioned tasks: e.g., the WESAD [29], CLAS [30], DEAP [31], MuSe [32], MDPSD [33], MMSD [27] and the new StressID [34].

A. OBJECTIVE AND CONTRIBUTION

The individual stress responses are both extremely non-specific [35] and affected by the relation between the subjects and the environment [2]. Therefore, considering merely the intrinsic stress, i.e., the one induced by the task itself, may be insufficient. This will result in a set of complications, especially if the CORT levels are not monitored: e.g., the investigated signals may or may not manifest evident changes during the task and if they are used to train an ML algorithm, errors in stress estimation could arise. Furthermore, not paying adequate attention to all the exogenous sources of noise affecting the collected signals may lead to misleading results and interpretations.

The above highlights the importance and necessity of a good experimental protocol. For this reason, the presented topical literature review aims to present approaches and solutions for inducing and assessing acute mental stress considering both psychological and technical insights. In the scientific literature, it is possible to find reviews on stress inducement and assessment with different levels of granularity and emphasis. Some explore the type of stressors to be presented during the experiments [36], [37], others focus mainly on the signals and features that can be used for stress analysis [4], [38], whereas others analyze the different algorithms for automatic stress detection [19], [24]. This review is intended to extend the existing literature by providing a practical approach, from participant assessment to physiological data collection, for implementing protocols to induce stress assessment in controlled conditions. To the best of our knowledge, this is the first work aiming to investigate this topic from two distinct perspectives: the psychological and engineering ones. We believe that the contribution of this paper, influenced by the interplay of authors with different backgrounds and expertise, could strengthen future studies by considering stress perception and the sources of inaccuracies that could comprise the data collection and analysis, thus limiting the confounding factors.

B. STRUCTURE OF THE WORK

The rest of the paper is structured as follows. Section II will provide a detailed presentation of the physiological events that occur in the presence of a stressful condition. Section III will interest stress-inducing approaches in controlled conditions, whereas Section IV will present quantitative tools (questionnaires) for stress assessment and the main sources of inaccuracy that could compromise the mostly adopted signals used for stress detection. Finally, Section V will be dedicated to the conclusion including some final remarks and a future perspective.

II. PHYSIOLOGICAL RESPONSE TO STRESSORS

Although the human stress response is extremely non-specific, the physiological one is well-defined. When a stressful event (also known as "stressor", i.e. exogenous or endogenous stimuli that evoke a stress response [4]) occurs, a cascade of hormonal and physiological changes manifests, involving both the hypothalamus-pituitary-adrenal axis (HPAA) [40], [41] and the sympathetic-adrenal medullary system (SAMS) [42], [43]. Both the HPAA and SAMS will be presented in this section. Note that in the subsection relative to the HPAA and SAM system we will refer to Fig. 1A and 1B, respectively.

A. HYPOTHALAMUS-PITUITARY-ADRENAL AXIS (HPAA)

When a stressful event occurs, its severity and emotional impact are assessed by the amygdala which stimulates the hypothalamus in case of threat. This results in the secretion of the corticotropin-releasing hormone (CRH) (1), which travels in the hypothalamic-pituitary portal system to the pituitary gland (2). As a consequence, the anterior pituitary secretes the adrenocorticotropic hormone (ACTH)(2) that flows in the bloodstream until reaching the adrenal glands (3).

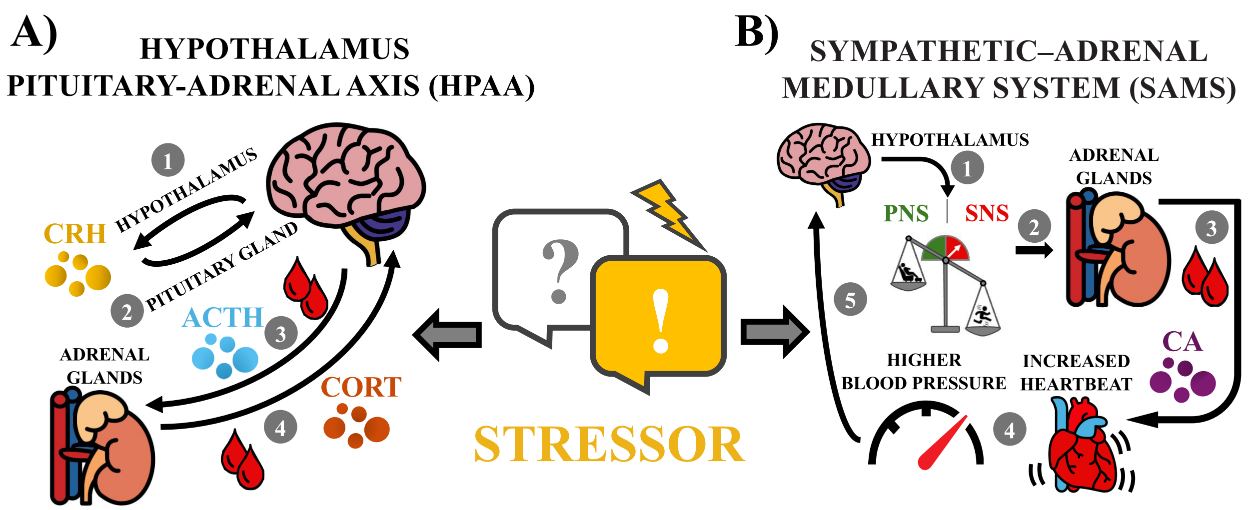

FIGURE 1. **The physiological response to a stressful event. (A) Activation of the hypothalamus-pituitary-adrenal axis (HPAA), resulting in cortisol secretion by the adrenal glands in the blood flow. (B) Activation of the sympathetic-adrenal-medullary system (SAMS), leading to the onset of the sympathetic nervous system (SNS) activity, i.e., the main component of the *"fight or flight"* response. As a consequence, physiological changes such as increased heartbeat and elevated blood pressure occur. For further details, the reader is referred to [10] where a comprehensive list of all the physiological changes that manifest after the SNS activation is presented.**

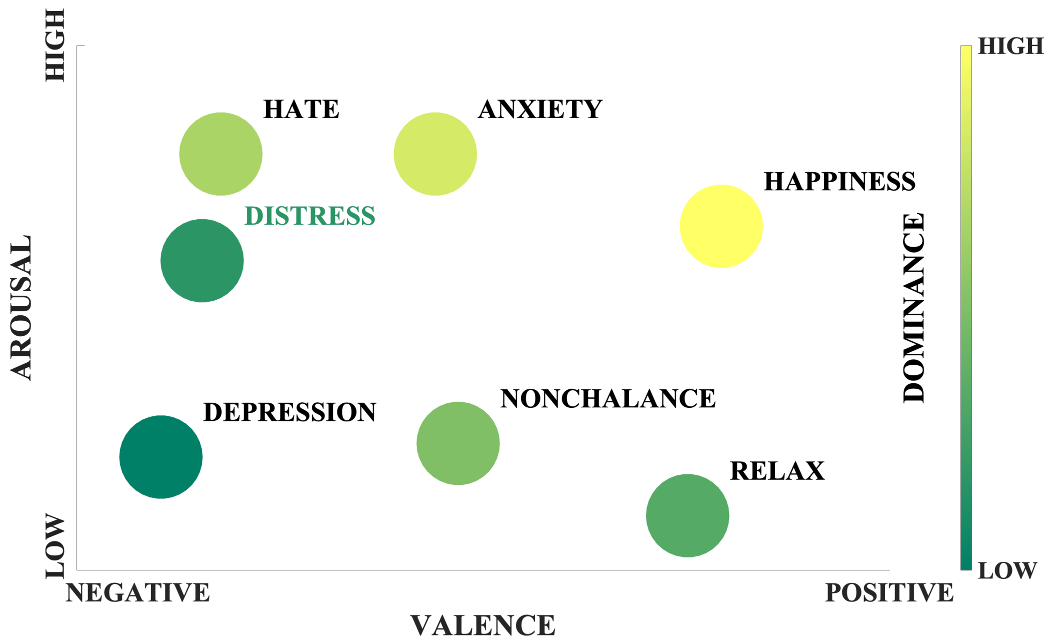

FIGURE 2. **Graphical representation of a set of emotions on the valence-arousal plane. The dominance dimension is considered through a colour map ranging from dark green (low-dominance) to yellow (high-dominance). Distress (or Stress) is located on the left part of the plane and it is characterized by a low-dominance. The mean values of the tables reported in [39] were considered to locate the emotions in this figure.**

Finally, the presence of ACTH induces the release of CORT (4), which is known to be the major stress hormone [44], in the bloodstream. In the acute phase, elevated CORT stimulates catabolic processes to provide the body with extra energy [45], involving the liver, muscles, adipose tissue, and pancreas [46].

B. SYMPATHETIC-ADRENAL-MEDULLARY SYSTEM (SAMS)

Similarly to what happens for the HPAA, the amygdala appraises an ongoing event as a threat and stimulates the hypothalamus (1), which then emphasises the activity of the sympathetic nervous system (SNS - antagonist of the parasympathetic nervous system, PSN) (2). The increased activity of the SNS induces the adrenal glands to secrete catecholamines (CA): adrenaline (epinephrine) and noradrenaline (norepinephrine), into the bloodstream (3). This results in an increase in different physiological parameters, such as the heartbeat and blood pressure (4). Notably, other physiological changes are also involved due to the activation of the sympathetic nervous system. For further details, the reader is referred to [10]. Similarly to what happens with the HPAA, the central nervous system continuously adapts the release of hormones in response to the appraisal of the situation (5).

In contrast to the analysis of HPAA activity, which requires either salivary, hair, urinary, plasma, serum or sweat fingernails samples [26] and can not be performed in real-time [27], the effects of SNS and PNS can be monitored through physiological signals continuously (see Section IV).

III. INDUCING ACUTE MENTAL STRESS

A. PSYCHOLOGICAL FUNDAMENTALS

Stress can be divided into two categories: eustress (related to positive events) and distress (associated with negative events) [4], the latter of interest in this paper. Since each emotion is characterized by three components: valence (ranging from positive to negative), arousal (varying from low to high) and dominance (ranging from low to high control on the events) [39], distress can be located on the left side of the valence-arousal plane, i.e., negative valence and high arousal. This condition is also associated with low dominance, as represented in Fig. 2, where other emotions are considered for comparison. Notably, the emotion's location on the arousal-valence plane is obtained considering the mean values reported in [39] tables. Despite the importance of distinguishing between eustress and distress, from now we will simply refer to "distress" with "stress".

From a psychological perspective, when facing a stressful condition, e.g. elicited through a cognitive task, two types of coping strategies may manifest. On one side we have active coping strategies, i.e., approaches including both thoughts and behaviours for managing internal and external demands perceived as stressful [47]. On the other hand, there are the strategies of passive coping, which are conditioned by feelings of helplessness and the need for external support [48].

Because of the subjective stress response, it is difficult to assume that an event will always elicit stress. Notably, individuals will perceive an event as stressful by considering both personal factors, i.e, the subjective perception in the environment [49], [50] and situational factors, i.e., appraisals of the subject with the environment [50], [51]. Therefore, both should be considered to make the event significant, resulting in a stressful response [2]. From now on, we will consider as stressful stimulus a generic cognitive task, better explained in the next section.

The process of generating acute mental stress can be graphically resumed in Fig. 3. The chosen task will induce intrinsic stress in the candidate and hopefully a stressful condition. Notably, stress may subtract cognitive resources and thus a reduction in performance is expected [52]. Based on the above, extrinsic stressors, i.e., due to external factors not related to the task, need to be considered to "boost" the stress perception. For this purpose, the factors (or pressure variables) proposed in [53] could be considered: i.e., audience (and the related factors), competition, reward and punishment, ego relevance and the probability of not being given another chance. In the following, strategies to implement the mentioned factors will be presented.

It has been demonstrated that the sex of the experimenters (or judges) enrolled in the investigation influences the stress perception of the volunteers [54]. In this regard, both men and women show higher levels of CORT when performing in front of judges of the opposite sex [55]. The audience's effect is not limited to the physical presence of people in the room. Indeed, video-recording the subject while performing the task affects the HPAA activation [56] and amplifies both the stress response and the feelings related to threat [57]. Furthermore, this approach would add pressure via at least two factors outlined in [53]: social comparison and social evaluation. If adopted, it is recommended to inform participants that they will be recorded during the experiment and a panel of judges will assess their non-verbal communication, as recommended in [58].

Setting up the experiment like a tournament, where the performances of the participants will be compared, could increase the CORT response, as demonstrated in [59]. The increased stress perception due to competition is linked to social-evaluation threats, that manifest when poor performances are evaluated by judges [60]. In the context of a social evaluation, feedback has a role in the experiments. Derogatory feedback elicits negative affect, i.e. characterized by unpleasant emotions, amplifying negative emotions [61]. In this regard, despite statistical differences in backward counting performances were not obtained, negative feedback induced a higher number of errors if compared to positive feedback in [62].

For what concerns the rewards and punishment factors, people receiving monetary rewards performed worse because of this pressure aspect [63], [64]. On the other hand, punishments could be set up with follow-up interviews that have been shown to create pressure and heighten task engagement in past research [65], even if the interviews never actually materialize. In this regard, it is worth noting that the use of deception in research requires careful consideration and strong justification from an ethical perspective.

Evaluative tasks are additional approaches for boosting stress perception, exploiting ego relevance pressure variables. Assessing the level of intelligence could be a good approach in this regard, especially when students are involved [53].

The mentioned factors could be taken together to boost stress perception in the chosen experiment, in addition to the possibility of not having a second chance to perform the task, similar to what happens in sportive scenarios [66].

In addition, the eight situational properties of stressors listed in [2] could be considered when artificially creating stress, which includes the concepts of novelty, event uncertainty, imminence, duration, temporal uncertainty, ambiguity and timing in relation to the life cycle of the user.

B. STRESSFUL TASKS

It is possible to induce acute mental stress with different strategies. In this regard, the reader is referred to [67] where a comprehensive table including the clinical and laboratory methods is reported. This section will focus on the common tasks used in the experimental protocols of the public databases mentioned (see Introduction). Moreover, suggestions for their implementation and improved versions will be discussed.

Regardless of the task chosen, a rest phase is usually observed [68]. This step is known as ''Baseline'' and permits subjects to acclimatize and relax before the task(s). There are several approaches for this purpose: let the volunteers relax by providing neutral reading material like magazines [31], [69], ask participants to watch a cross on the screen [31], or relaxing videos [70], with particular attention to the ones involving natural settings, which are considered the most relaxing ones [71]. The temporal duration of the baseline condition is not fixed, but a 15-minute adaptation period is suggested [68].

The Mental Arithmetic Task (MAT) is certainly one of the most adopted cognitive tasks. It is possible to implement it in different ways: e.g, iterative subtraction of a fixed number from a starting one [72], additions [73] or with multiple operations (additions, subtractions, multiplications and divisions) [74]. The temporal length for this task is preferred to be 5 minutes [58]. To increase the stress perception, it is possible to modify the task slightly; for example, by asking the participant to answer the highest number of questions possible [73] or doing the calculation out loud [75]. The MAT is the basis of other cognitive tasks like the Montreal Imaging Stress Task [76]. The latter requires the participants to provide the result of calculus within a certain amount of time, which may increase or decrease depending on whether the user is performing well. To enhance the stress perception, the user's performance is shown in the interface and compared with the previously tested subjects [76]. The Paced Auditory Serial Addition Task (PASAT) is also based on the MAT and can be used as a stressor in controlled conditions [77]. The test is structured as follows: participants are asked to sum adjacent numbers of an array, before the presentation of the new series [78]. The difficulty can be increased by reducing the available time for solving the mathematical operations to as little as 1.2 seconds [78]. Furthermore, by enriching the PASAT with emotional, acoustic and motivational stressors (together with decreased answer time) the Mannheim Multicomponent Stress Test can be obtained [79]. The latter is a solid approach for stress induction in controlled conditions [67], [80] and exploits images with high arousal and negative valence, white noises and monetary punishments as emotional, acoustic and motivational stressors [80], [81]. To increase the complexity of the task, one of the previous factors could be modified (e.g., the cost of each error).

Another widely used cognitive task to induce stress in controlled conditions is the Stroop Colour Word Task (SCWT) [82], which involves identifying the colour of a word shown on the screen. The stimuli in this test can be divided into three groups: neutral (word written in black - used to assess readability), congruent (word written with the same target colour) and incongruent (word written with an ink different from the target colour), with the latter considered the most challenging [36], requiring more time to be completed [83]. Also for this task, reducing the available time to provide the answer will induce pressure on the participants.

The Speech Task (ST) is yet another extensively used method. Participants are asked to prepare a speech in a specific amount of time (ranging from 5 to 10 minutes [84], [85]) and present it in front of a committee without notes. The topics of the speech are variegated. However, the defence against the charge of shoplifting [84] and job interviews [85] are the most adopted [58]. To increase the stress perception during this task, different approaches could be adopted. Preparing a speech under challenging conditions, such as with background noise [86] or within a short time frame (e.g., 3 minutes [56]), are just two examples. Another effective approach is practicing in front of a camera and microphone [87]. Note that the combination of the previous devices is supposed to induce higher grades of social evaluation [88]. Variations of the ST are the Sing-A-Song Stress Test [89] and simple singing stress procedure [90], both demanding the subjects to sing a song rather than to speak.

The above tasks exploit active coping strategies, which are defined in Section III. It is worth noting that it is possible to elicit stress also with other approaches, based on passive coping strategies. Viewing images with high arousal and negative valence (e.g., the ones available in the International Affective Picture System (IAPS) [91] and Open Affective Standardized Image Set (OASIS) [92] datasets), videos (e.g., spider-related videos to individuals who are afraid of the spiders [93]) and performing the cold-pressure test (i.e., immersing one hand in a bucket of cold water for a certain amount of time) are common examples that fall into this category.

Additionally, there are tests that combine some of the previously mentioned tasks. The Trier Social Stress Test [56], [85], which exploits both the ST and the MAT, is certainly the best example, being the most used approach for stress induction in controlled conditions [58]. Different variations of this test exist, differing from the specific samples (e.g., children [94] or adolescents [95]), contexts (e.g., in person [85] or online [96]) and condition (e.g., in a virtual reality scenario [97]). Another example is the Maastricht Acute Stress Test [98], which exploits the MAT and the

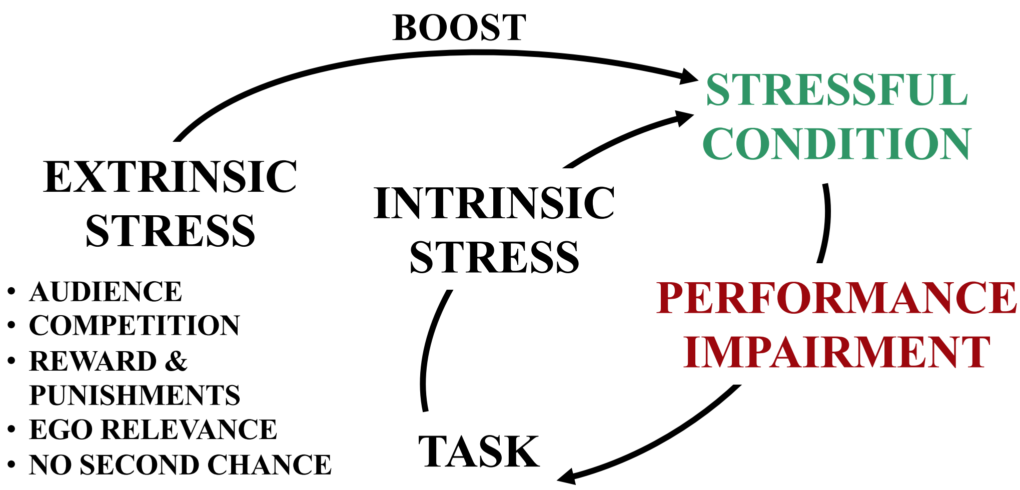

FIGURE 3. Generation of a stressful condition through the combination of both intrinsic and extrinsic stressors. The first pertains to task execution, whereas the others involve external factors unrelated to the task. To elicit an extrinsic stress response, the factors presented in [53] could be considered: i.e., audience, competition, reward and punishment, ego relevance and the probability of not having a second chance. The combination of both intrinsic and extrinsic stressors leads to a higher perception of stress by the users (which could subtract cognitive resources [52]) and thus a reduction in performance.

cold-pressure test to enhance the stress perception of the users.

IV. ASSESSING ACUTE MENTAL STRESS

A. PARTICIPANT ASSESSMENT

Although being superficial or even missing in some studies, an in-depth analysis of the volunteers taking part in the experiment is fundamental.

When designing an experimental protocol, regardless of the aim of the study, gender-balanced samples are needed. Since differences in responses to stressors exist among sexes [106], grouping data from both males and females may mask the effects due to sex [107]. Moreover, particular attention should be taken when women are enrolled; indeed, their CORT level depends on the luteal phases and may affect emotional and cognitive responses when performing a stressful task [108], [109].

As presented in Section III, it is possible to induce acute mental stress through cognitive tasks. However, it is worth noting that humans react differently based on age and education [110]. Therefore, information relative to both age and education level should be provided, defining specific inclusion criteria when enrolling participants.

Despite existing controversial results about the importance of body weight on neuroendocrine system functioning [111], we believe it is an important feature to consider. Indeed, excessive weight may compromise both optical measurements, e.g., the photoplethysmogram (PPG) [112], and the electrical ones like the electromyogram (EMG) [113]. For further details, the reader is referred to Section IV.

Pharmacological treatments impact behavioural and physiological response (e.g., CORT production) [58]. Therefore, supplements and non-essential drugs should not be consumed in the 24 hours before the experiments. Mixed findings exist on tobacco consumption [58], whereas the assumption of caffeine seems to affect the circadian rhythms [114] and the brain activity [115]. As remarked for the assumption of non-essential drugs, it would be preferable to avoid their consumption the day before the experiments.

The moment of the day when the experiment is performed and the sleeping hours may impact the stress perception. Because of the circadian cycles, which influence the levels of CORT [116], performing a stressful task at different times of the day may lead to different outcomes. For this reason, it is suggested to experiment in the same phase of the day (either in the morning or afternoon) for all the participants. Working night shifts [117] and sleep shifts [118] impact the HPAA activity and, as a consequence, the CORT secretion. This information is missing in most of the studies and particular attention should be taken during participant screening, especially if involving undergraduate students [58], being more prone to manifest chronically limited sleep [119]. Even though the optimal number of sleep hours is strongly subjective and depends on the age [120], we suggest participants sleep 7 to 9 hours before the experiments.

Personality traits (e.g., neuroticism, extroversion), cultural identity and the history of traumatic events are key details to consider when recruiting volunteers for an experiment aiming to explore their stress response. Personality traits can be studied through the multidimensional personality questionnaire, which examines three orthogonal higher-order factors: Positive Emotionality (PEM), Negative Emotionality (NEM), and Constraint (CON) [121]. In this regard, it has

TABLE 1. List of questionnaires mostly used in experimental protocols. The table includes the target component to be analysed, the number of questions, values for each mental condition (e.g., low, moderate, or high stress), and the phase of the experiment when the questionnaire can be proposed (e.g., before, between and after the task(s)).

Questionnaire	Questions	Values	Ranges	Phase	Ref.
Perceived Stress Scale (PSS)	10 (4 - 14)	$0 \rightarrowtail 4$	**Low**: [0 - 13] **Moderate**: [14 - 26] **High**: [27 - 40]	**Before**	[99]
State and Trait Anxiety Inventory (STAI)	40	$0 \rightarrowtail 4$	**Low**: [20 - 37] **Moderate**: [38 - 44] **High**: [45 - 80]	**Before** **After**	[100]
Visual Analogue Scale (VAS)	1	$0 \rightarrowtail 10$	**Low**: [0 - 5[**Moderate**: [5 - 8.2] **High**:]8.2 - 10]	**Before** **Between** **After**	[101]
Self-Assessment Manikin (SAM)	3	$1 \rightarrowtail 5$	**Scale** / **Valence** / **Arousal** / **Dominance** 5 / Pleasant / Excited / Dependent 4 / Pleased / Wide-awake / Powerlessness 3 / Neutral / Neutral / Neutral 2 / Unsatisfied / Dull / Powerful 1 / Unpleasant / Calm / Independent	**Before** **Between** **After**	[102]
NASA Task Load Index (NASA-TLX)	6	$0 \rightarrowtail 100$	**Low**: [0 - 9] **Medium**: [10 - 29] **Somewhat high**: [30 - 49] **High**: [50 - 79] **Very high**: [80 - 100]	**Between** **After**	[103]
Positive and Negative Affect Schedule (PANAS)	20 (10)	$1 \rightarrowtail 5$	**Positive or Negative Affect based on Total Score**	**Before** **After**	[104] [105]

been demonstrated that participants with high NEM exhibited great emotional stress [122] when performing the TSST. In terms of cultural identity, individuals from different cultural backgrounds (e.g. Eastern vs. Western cultures) [123], [124], [125] perceive and evaluate stressors differently, as well as volunteers with a history of traumatic events. Individuals with Post-Traumatic Stress Disorder (PTSD) often exhibit heightened stress when faced with unexpected stimuli [126]. All the mentioned factors can profoundly shape how individuals perceive and cope with stressful events, ultimately influencing experimental outcomes, thus justifying the necessity to consider those details during the participant assessment.

All the above should be considered in the preliminary phase of the experiment together with the exclusion of the participants who performed vigorous exercise in the previous 24 hours and those suffering from severe physical and mental health. Overall, to conduct an in-depth analysis of the participant, we refer the reader to [58], where a template considering most of the mentioned factors for this purpose can be found.

B. SELF-REPORT ASSESSMENT

Once an objective assessment of the participants has been conducted, questionnaires can be adopted to assess subjective experience [127]. Altogether, they differ in target (e.g., stress, anxiety, emotions, mood, mental workload), number of questions, and time required to be filled in. To have a general overview of the participants, at least two questionnaires should be considered during the experimental protocol [58]. In this section, some of the most used will be discussed, providing also details about the number of questions, values for each mental condition (e.g., low, moderate, or high stress), preferred phase to be present them during the experiment (e.g., before, during or after the task). The discussion is summarized in Table 1.

- **Perceived Stress Scale (PSS)**: introduced in 1988 [128], usually exploits 10 questions with the possibility to extend and reduce them to 14 and 4, respectively. It can be used before the beginning of the experiment. The resulting score ranges from 0 to 100 when 10 questions are considered and high levels of stress are usually associated with values between 27 and 40 [100]. Despite being useful and commonly used, it is important to underline that the PSS assesses also a range of stress-related concepts including appraisals, coping, and emotions like anxiety.
- **State and Trait Anxiety Inventory (STAI)**: released in 1983, includes a set of 40 questions, divided into state and trait anxiety scales, to measure the state of anxiety [129]. Whereas the trait anxiety scale is useful to give a feel for the general level of anxiety a participant is likely bringing into the experiment (e.g., like a confounding factor), the state anxiety scale is useful to indicate if anxiety changes as a result of (i.e., before and after) an experimental manipulation such as a

high-pressure or stressful task. It can be presented to the volunteers before and after the task and the resulting score ranges from 0 to 80. High levels of anxiety are related to scores between 45 and 80 [100].

- **Visual Analogue Scale (VAS)**: used to assess the symptom severity related to a disease, it is a common tool in the clinical practice [130]. In the context of stress measurement, a coloured scale (codified in 10-points) is preferred [131]. It is extremely rapid and misunderstandings can be avoided [132]. Because of its rapidity, it can be presented to the participants not only before and after the experiment but also between tasks. The resulting score ranges from 0 to 10 and a score between 8.2 and 10 can be linked to a high-stress perception, according to [101]. Similarly to the visual analogue scale, other questionnaires can benefit from its rapidity. In this regard, the reader is referred to the stress [133] and anxiety [134] thermometers and the immediate anxiety measurement scale [135], just to mention a few.
- **Positive and Negative Affect Schedule (PANAS)**: introduced in 1988 [104], it consists of 20 questions, 10 measuring the positive affect (e.g., interested, active) and the remaining the negative one (e.g., irritable, nervous) [136]. With this questionnaire, it is possible to quantify mood and emotions. The prevalent affect (positive or negative) is the one having the highest score. Despite the existence of a shorter version with only 10 questions [105], it is not as rapid as the visual analogue scale. For this reason, it is preferable to present it before and after the experiment.
- **Self-Assessment Manikin (SAM)**: it is a non-verbal tool to assess emotions and was introduced in 1994 [137]. It explores three principal dimensions: valence, arousal and dominance. This questionnaire is rapid to be filled out and can be presented to the participants not only before and after the experiment but also between tasks. Notably, stress conditions are associated with negative valence, high arousal and low dominance (see Fig. 2).
- **NASA-Task Load Index (NASA-TLX)**: it is a questionnaire used to assess both the physical and mental workload of a task, exploring six dimensions: mental, physical, and temporal demand, along with performance, effort, and frustration [138]. The resulting score ranges from 0 to 100. A score between 50 and 79 is associated with a high mental workload, whereas if the total score falls between 80 and 100, it can be classified as very high demand [103]. It is possible to consider variants of the presented questionnaire: e.g., the SURG-Task Load Index and the SIM-Task Load Index. The first explores mental, physical and temporal demands together with task complexity, situational stress and distractions [139], whereas the latter includes additional factors like perceptual strain, task control and presence [140], making it suitable for demands of simulated environments in virtual reality.

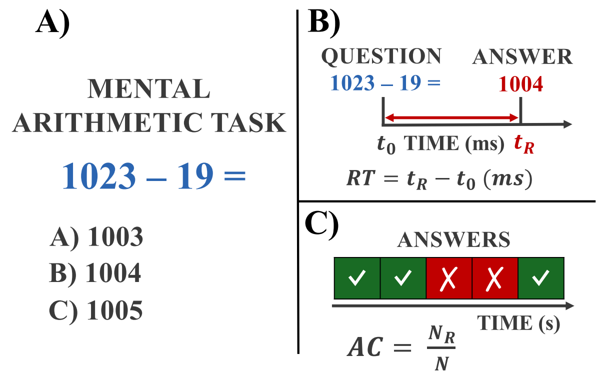

FIGURE 4. A) Example of question for the mental arithmetic task. (B) Reaction time (RT) refers to the time between the question's display (t_0) and the moment when the user provides the answer (t_R). (C) Accuracy considers the number of corrected answers (N_R) over the total number (N).

Collecting data through questionnaires is certainly an immediate and inexpensive strategy. However, attention should be taken. Because of a strong subjective component and social desirability bias (i.e., the tendency to over-report more desirable qualities and under-report socially undesirable actions [141]), they can not be completely reliable. This results in a set of limitations, including biases in the estimations [34]. In addition, questionnaires along with the level of CORT and many other biological markers cannot be collected continuously [28].

C. PERFORMANCES

With particular attention to the MAT and the SCWT, the number of correct answers and reaction times (RT) need to be monitored for the whole duration of the tasks [142]. RT can be defined as the time between the onset of the question and the moment in which the user provides the answer, whereas the number of correct answers over the total number of questions in a specific temporal interval is known as accuracy (AC) (see Fig. 4). Performing a task in acute stress conditions will impact both RTs and ACs. More specifically, an increase of RT [143], [144] and a decrease in AC [145], [146] is expected.

By combining the mean RT of the correct answers and AC the inverse efficiency score (IES) can be obtained:

$$IES = \frac{RT}{AC} \tag{1}$$

It is usually measured in milliseconds and takes into account both the velocity and precision while performing a specific task. Notably, it is possible to measure this index also for a limited number of questions to compare the performances over the tasks. In this case, the Equation (1) can be adapted as follows:

$$IES = \frac{\frac{\sum_{k=1}^{K} RT_k \cdot a_k}{(\sum_{k=1}^{K} a_k)}}{\frac{(\sum_{k=1}^{K} a_k)}{K}} = K \frac{\sum_{k=1}^{K} RT_k \cdot a_k}{(\sum_{k=1}^{K} a_k)^2}$$

$$a_k = \begin{cases} 1, & N_k = 1, \\ 0, & \text{elsewhere} \end{cases} \tag{2}$$

where K is the number of questions in the interval, N_k is the answer to the k^{th} question, RT_k is the reaction time at the k position and a_k is a binary coefficient that can be either 1 if the user answers correctly to the question N_k or 0. Note that the numerator takes into account the mean reaction time of the correct answers, whereas the denominator is the accuracy. However, attention must be paid to the estimation of this parameter. Indeed, the high number of errors will bias the IES estimation, also making it unstable because only a few RT values are considered. For this reason, as suggested in [147], IES should be estimated when there is a strong correlation between RT and AC, considering the cases when the number of errors is low.

If the ST is considered, features relative to the performance could be estimated as well. In this regard, metrics like the number of words per minute and word productivity (i.e., the ratio between the number of effective words and the total number), could be assessed. However, pauses (i.e., duration of time between words) and their number seem to be the ones mostly affected by stress conditions [84].

D. PHYSIOLOGICAL SIGNALS

As a direct consequence of the HPAA and SNS activation, changes in subjects' homeostasis manifest, most of which can be detected through physiological signals.

Given the number of reviews focusing mainly on the signals and relative features that could be used for stress detection [4], [24], which the reader can rely on for further details, this section will have a practical connotation. A comprehensive list of the investigated signals is presented in Table 2, reporting the bandwidth of each signal (with the suggested sampling frequency), the main exogenous sources of noises and the suggested approaches to mitigate (when possible) their effect. Note that the software approaches to mitigate the effects of external sources of noise will not be considered.

Before analyzing the signals for stress detection, it is worth noting that data are sampled at a constant rate, and their Fourier transform results in periodic repetitions (aliases) of the original signal's frequency components. To avoid aliasing, a low-pass filter with a cutoff below the Nyquist limit (half the sampling rate) should be used before sampling to remove higher frequencies. This ensures accurate signal representation without high-frequency distortion [148]. With particular reference to the sampling frequency, Table 2 reports the common sampling frequencies for the investigated signals that exceed 3 times the maximum bandwidth of the signal, to compensate for the non-idealities of the anti-aliasing filters, which are part of the acquisition chain.

Heart rate variability (HRV) is widely adopted for stress detection [13], [14], [15]. However, attention should be taken when estimating it. Formally speaking, we are dealing with HRV only if the processed signal is the electrocardiogram (ECG), whose frequency band is mostly between 0.05 and 100.0 Hz [178]. On the other hand, if the photoplethysmogram (PPG) is considered, we are addressing the pulse rate variability (PRV). Note that the frequency band of PPG is usually between 0.01 and 10.0 Hz [179] and the suggested sample frequencies for the estimation of HRV and PRV are 250.0 Hz [180] and 50.0 Hz [153], respectively. Although a solid consensus regarding the applicability of the PRV as a direct substitute for the HRV is missing [181], [182], some studies demonstrated a correlation between the HRV features with the ones estimated through PPG, but only in rest conditions [181], [183], [184]. Nonetheless, it is worth pointing out that a high correlation does not imply interchangeability. Two main factors that do not make HRV and PRV interchangeable are the blood pressure [182] and the pulse-transit-time (i.e., the time required for the arterial pulse wave to travel to the sensor, PTT) [185], [186], which depends on the blood pressure as well [187]. Notably, since there is also an unknown time related to aortic valve opening, measuring the vascular transit time rather than the PTT is preferred [187].

For what concerns data acquisition, both ECG and PPG are susceptible to motion artifacts [188], [189], [190] that, if not attenuated, could compromise the signal-to-noise ratio (SNR). Different environmental factors affect the quality of the measurements: the ECG is influenced by the skin-electrode interface and by electromagnetic interference (EMI) [150], [191], whereas PPG is affected by the environmental light and temperature [112], [150]. Additionally, the skin tone [152] and fat thickness influence the quality of the PPG [112], [152]. To reduce the effect of the previous sources of inaccuracy, skin preparation [151], short cables [150], wet electrodes [151] and high common-mode-rejection-ratio (CMRR) amplifiers should be adopted for the ECG. On the other hand, if PPG is considered, it is possible to partially mitigate the previously highlighted sources. Keeping both the room temperature and humidity constant is certainly a good approach, along with the optical shielding [112]. Furthermore, to enhance the quality of the PPG, locations with high skin perfusion should be considered [192], e.g., the finger and the ear [193]. The effect of the skin tone could be slightly attenuated by increasing the wavelength of the light source [112]. However, it is not always possible to change this parameter. Devices collecting PPG (e.g. wristbands) are perfectly suited for long-time measurements, whereas ECG devices exploiting wet electrodes are sub-optimal since skin-irritation may occur [194]. Capacitive electrodes could be adopted as alternative [195], e.g., embedded into a

TABLE 2. Table including the widely adopted signals for stress detection. Each line of the table includes the type of signal, its frequency band (with commonly adopted sampling frequency), the main sources of inaccuracy (e.g., interference and noise) and the recommended actions to mitigate their effects.

Signal	Frequency Band (Sampling Frequency)	Exogenous Sources of Noise	Suggested Acquisition conditions	Ref.
Electrocardiogram	0.05 $\rightarrowtail$ 100.0 Hz (500.0 Hz)	• Motion artifacts, • Electromagnetic interference, • Skin-electrode interface.	• Skin preparation, • Short and not stretched cables, with wet electrodes, • High CMRR amplifiers.	[149] [150] [151]
Photoplethysmogram	0.01 $\rightarrowtail$ 10.0 Hz (50.0 Hz)	• Motion artifacts, • Ambient factors (temperature, humidity, light), • Skin tone, • Sensor location.	• Controlled temperature, humidity and light, • Optical shielding, • Site with high vascularization (e.g, finger and ear), • Wavelength selection (if possible).	[112] [152] [153]
Electrodermal Activity	0.03 $\rightarrowtail$ 0.5 Hz (4.0 Hz)	• Motion artifacts, • Ambient factors (temperature and humidity), • Skin-electrode interface.	• Controlled temperature and humidity, • Pre-gelled electrodes, • EDA conductive gel (ECG and EEG ones are not suitable), • Appropriate sites for measure: e.g., finger, foot and shoulder.	[154] [155] [156] [157]
Electromyogram	1.0 $\rightarrowtail$ 500.0 Hz (2000.0 Hz)	• Motion artifacts, • Electromagnetic interference, • Skin-electrode interface.	• Skin preparation, • Short and not stretched cables, with wet electrodes • High CMRR amplifiers, • Differential measurements. • Electrodes with fixed inter-electrode distance.	[158] [159]
Electroencephalogram	0.1 $\rightarrowtail$ 40.0 Hz (500.0 Hz)	• Motion artifacts, • Electromagnetic interference, • Skin-electrode interface,	• Short and unstretched cables, with wet electrodes • High CMRR amplifiers, • Differential measurements.	[150] [160]
Surface Temperature	0.001 $\rightarrowtail$ 0.1 Hz (1.0 Hz)	• Skin tone, • Skin to sensor distance, • Ambient factors (e.g,, wind, temperature, wind's speed).	• Fixed distance between the skin and the sensor, • Controlled ambient temperature.	[161] [162]
Respiration	0.1 $\rightarrowtail$ 1.0 Hz (4.0 Hz)	• Modulation of the cardiac activity: see *Electrocardiogram* and *Photoplethysmogram*. • Chest-wall movements: -- Motion artifacts, -- Temperature and humidity (for piezoelectric sensors).	• Modulation of the cardiac activity: see *Electrocardiogram* and *Photoplethysmogram*. • Chest-wall movements: -- Sensor in a chest strap, -- Protection of the piezoelectric sensor with shielding materials.	[163] [164] [165] [166] [167]

TABLE 2. *(Continued.)* Table including the widely adopted signals for stress detection. Each line of the table includes the type of signal, its frequency band (with commonly adopted sampling frequency), the main sources of inaccuracy (e.g., interference and noise) and the recommended actions to mitigate their effects.

Signal	Frequency Band (Sampling Frequency)	Exogenous Sources of Noise	Suggested Acquisition conditions	Ref.
Pupillogram	0.2 ⤚ 4.0 Hz (20.0 Hz)	• Motion artifacts, • Ambient factors (light, temperature).	• Fixed eye-camera distance, • Controlled ambient temperature (20-23°C), • Controlled ambient light.	[168] [169] [170]
Speech	0.1 ⤚ 20.0 kHz (44.1 kHz)	• Environmental noise, • Recording distance, • Type of microphone.	• Condenser microphone with pop-filters, • Fixed distance from microphone (unidirectional), • Mono recording with uncompressed format (WAV), • Soundproof room (if possible).	[171] [172] [173] [172]
Inertial Measurement Units	0 ⤚ 15.0 Hz (50.0 Hz)	• Ambient factors (temperature, humidity, light), • Vibrations, • Electromagnetic interference.	• Controlled ambient, temperature and humidity, • Manual calibration, • Shielded wires.	[174] [175] [176] [177]

garment [195], [196]. However, this solution is more prone to motion artifacts [197].

The electrodermal activity (EDA) is a well-established for stress detection [16], [17], [18] and has two components [198]: the skin conductance level (SCL), and the skin conductance response (SCR), also known as tonic and phasic components, respectively. Of the two, the SCR is one of the major since it is an indirect index of SNS activation [199]. The frequency components of this signal fall between 0.03 Hz to 0.5 Hz, [156], [157], with most of the power between 0.045 and 0.15 Hz [156]. For what concerns the sampling frequency, a minimum of 2.0 Hz is suggested [200]. Data acquisition is mostly influenced by motion artifacts [201] and environmental conditions like humidity and temperature [18], [155], [202], the latter mostly affecting the tonic component rather than the phasic one [202]. It is preferable to monitor both temperature and humidity and to use pre-gelled electrodes, with EDA conductive gel, for data acquisition [155]. To enhance the quality of the measure three areas are suggested in [203]: fingers, foot and shoulders. Overall, the measurement is not intrusive, it requires a little time to be set up and, if included in wearable devices, is meant for long-term measurements [204].

The surface electromyogram (EMG) and the surface electroencephalogram (EEG) are electrical measurements, relative to muscle contraction and neural activity, respectively. Both the EEG [23], [205], [206] and the EMG [22], [207], [208] are used as signals for stress detection. The frequency band of the EMG is between 1.0 and 500.0 Hz and the most used sample frequency is around 2.0 kHz [158]. On the other hand, the EEG has a frequency band between 0.1 and 40.0 Hz and a good sample frequency value could be around 500 Hz. Electromagnetic interferences can influence both the EMG and EEG [209], [210], which can be reduced using amplifiers with appropriate CMRR and differential measurements. In this regard, particular attention is necessary when considering differential measurement. Indeed, impedance mismatches due to the skin-electrode interface may vanish the rejection of the common mode. Short and not stretched cables in combination with wet electrodes should be used to enhance the SNR [158], [159] and reduce the effect of motion artifacts. Note that both are susceptible to that artifact. This is true, especially for the EEG [211], because of the strong overlap between the frequencies of the motion artifacts and the brain rhythms. Regarding the time for the set-up, the EEG is certainly the most time-consuming. On the other hand, the time required for the EMG set-up increases with the number of electrodes [158].

Surface body temperature (SBT) is another physiological signal that could be considered for our purpose [212], [213], [214]. Indeed, to ensure homeothermy, the SNS induces peripheral vasoconstriction in case of acute stress [215] and consequently a reduction of the surface temperature [216].

SBT oscillates with frequencies ranging from 0.001 to 0.1 Hz [162] and a sampling frequency of 1.0 Hz is certainly more than enough for data collection. If measured through non-contact infrared thermometers, surface body temperature is affected by both environmental factors (e.g., ambient temperature, humidity, wind, heat sources) and acquisition conditions (e.g., distance and angle [161], [217]). Similarly to the PPG, SBT is influenced by the skin tone [161]. To overcome some of the previous aspects, it is recommended to monitor the environmental factors and guarantee a fixed distance between the sensor and the skin during the experiment.

The increased activity of the SNS impacts the respiration of the subjects. Since the respiration signal has a frequency band between 0.1 and 1.0 Hz [163], [167], a sampling frequency of 4.0 Hz is recommended. Despite the numerous contact-based methods for estimating the respiration signal [166], this section will focus on methods based on chest-wall movements and the modulation of cardiac activity. The first exploits sensors placed on the chest wall (e.g., resistive, capacitive, inductive and fibre optic sensors [166]), whereas the latter the ECG [164] and PPG [165]. If the respiration signal is estimated through the heart-related signals, the problems relative to ECG and PPG will be inherited. On the other hand, if the chest-wall movements are considered, the main source of inaccuracy is related to sensor movements (i.e., motion artifact), not related to the chest. For this reason, it is necessary to ensure good contact between the sensor and the skin, which could be embedded into a chest strap. The environmental conditions do not influence significantly the sensors used to analyze the chest-wall movements, except for the strain gauges based on piezoelectric sensors, which are sensitive to temperature and humidity changes [166], leading to output drifts. For this reason, it is recommended to cover the sensor with protective materials [166]. Overall, both the presented approaches show low set-up times and the possibility for long-term acquisitions.

The SNS influences also pupil size [218], [219] and studies demonstrated the efficacy of considering the pupil diameter for stress analysis and assessment [220], [221]. Its frequency range oscillates between 0.01 and 5.0 Hz [222] and the suggested sampling frequency is 20.0 Hz [168], [169]. Concerning the exogenous sources of noises that could compromise the measurement, motion artifacts and ambient conditions (i.e., light and temperature) are certainly the most influential [169], [223]. Thereby, experiments should not be performed in very dark or very bright environments [168], maintaining the distance between the eyes and the camera fixed (to reduce the head's movement) and monitoring both the background and room illuminance during the experiments [169], [170]. Because of SNC activation in cold environments, the suggested temperature for the experiment is between 20 and 23 °C [169].

As introduced in Section IV, changes in the speech manifest because of stressful situations. For this reason, speech analysis can be exploited for stress assessment [224], [225]. The frequency range of this signal is mostly between 0.1 and 8.0 kHz. Although higher frequencies may manifest during data acquisition, they are usually exploited for user localization [172]. The main sources of exogenous noise during the data collection are related to the environmental noise, the distance from the microphone and the type of device used [173]. To mitigate their effect, it is recommended to use condenser microphones with pop-filter, a set-up for keeping fixed the distance from the microphone and use a mono recording with uncompressed format (e.g., Waveform Audio File, WAV) [173].

Finally, data from the inertial measurement units (IMU) can considered as well. In most cases, accelerometer data is coupled with other signals [226], [227], [228], whereas sometimes they are the only signal considered [88], [229]. The maximum bandwidth of the human body movements falls between 10.0 and 15.0 Hz according to [175]. For this paper, the focus will be on IMU based on micro-electronic-mechanical systems (MEMS). The main external factors compromising their output are temperature changes, humidity, vibrations, shocks and electromagnetic fields [174], [176]. For this reason, the environmental condition should be monitored. If low-cost IMUs are used, usually coming with smartwatches or smartphones, additional systematic errors arise. In this regard, to mitigate both errors and possible drift recalibrations can be considered, also without external devices [177]. On the other hand, to mitigate EMI's effect shielded wires and cases should be used. Overall, they require little time to be set and can be leveraged for long-term acquisitions if embedded in either wristbands or smartwatches.

E. DATA ANALYSIS

The experimental protocol pipeline can be graphically summarized in Fig. 5, which is divided into four phases: the first implies the analysis of the participants and the self-assessment reports (the second one to be presented to the participants after the experiment's conclusion); the second phase pertains to the baseline condition in which the participants acclimatize and the physiological signals begin to be collected; the third involves again a self-assessment report to be filled out before the onset of the stressful event; and the final one, relative to the stress-inducing task. The types of questionnaires used may vary across different phases of the assessment. For instance, the initial self-report assessment might include questionnaires like the PSS, STAI, VAS, and PANAS, whereas the second assessment could utilize a similar set, excluding the PSS. In contrast, the final assessment may incorporate questionnaires like the STAI, PANAS, SAM, VAS, and NASA-TLX.

For ease of understanding, Fig. 5B and 5D show four physiological signals, commonly adopted for stress assessment, with one feature each. Data come from the WESAD database [29] and are relative to subject S14. The ECG was collected using a standard three-point measurement, the EMG

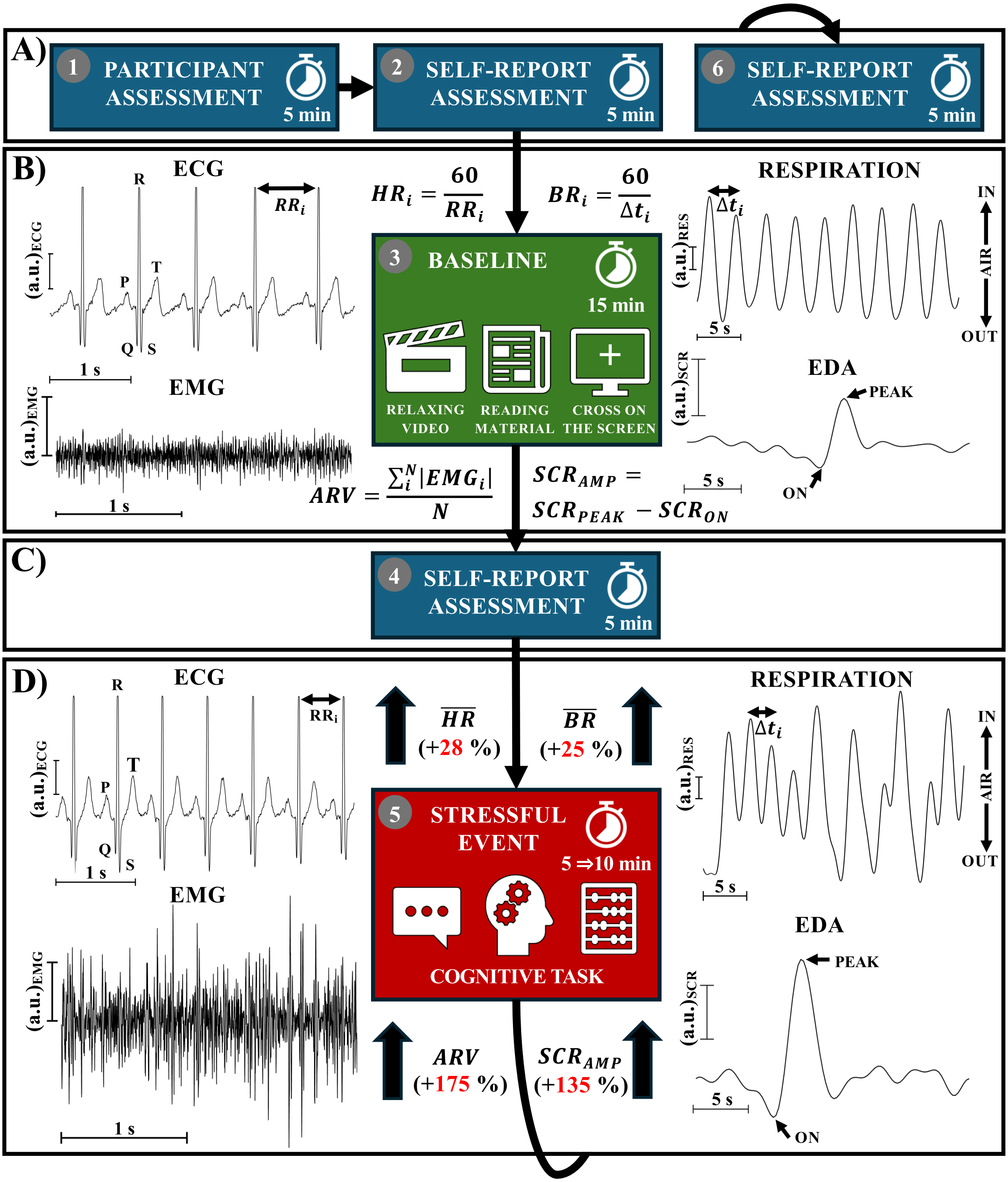

FIGURE 5. Graphical representation of the experimental protocol pipeline in four phases: A) Participant and self-report assessment, the latter repeated twice: at the beginning and end of the experiment B) The baseline condition, which is meant to let the participants acclimatize. Reading magazines, relaxing videos or looking at a cross on the screen are common approaches adopted in this phase, while physiological signals are recorded. In this example, signals like the ECG, EMG, respiration and EDA are reported. The experimental data come from the WESAD database [29], subject S14. The reported signals are treated as follows: the ECG is filtered between 100 mHz and 100 Hz, the EMG is high-pass filtered at 10 Hz (to mitigate the motion artifacts), the respiratory signal is filtered between 100 mHz and 1 Hz, whereas for the EDA signal, the onset and peak of SCR are obtained with the Neurokit2 toolkit (function *eda_process* with method *"neurokit"*). Features like the heart rate (HR), breath rate (BR), average rectified value (ARV) of the EMG, and SCR amplitude are reported as examples. C) Self-assessment report, before the onset of the stressful situation, which aims to elevate psychological pressure before they start the task. D) Stressful event induced through a cognitive task. Changes in the physiological signals occur e.g., elevated values of HR and BR (average) and increased ARV and SCR amplitude. The blocks are numbered from 1 to 6, and an approximate each step's duration is reported.

of the trapezius, which is commonly used in experiments for stress induction and assessment, was collected in a bipolar mode, the respiration signal was captured using an inductive plethysmographic sensor, and the EDA was measured on the subjects' wrists. All signals were recorded with the RespiBAN at a sampling frequency of 700 Hz, except for the EDA signal, which was collected at a sampling frequency of 4 Hz using the Empatica E4. The baseline condition allows for a comparison of the features and signals with the stress condition, which in this case was induced using the TSST. The signals were processed with MATLAB® (version 2024a) and Python™ (version 3.11.3). The analysis of the EDA was conducted with the Neurokit2 toolkit (version 0.2.7).

In this example, the physiological signals and their respective features exhibit changes in response to stress conditions, e.g., higher average values of heart rate (HR) and breathing rate (BR), higher amplitude of the EMG signal (estimated through the average rectified value, ARV), and higher amplitude of the tonic component of the EDA signal, compared to the baseline condition. Note that this aligns with the findings of [4].

However, these are not the only features used to assess the stress response of the candidates; many others may be taken into consideration. In this regard, we refer the reader to a review of the signals and features that are related to stress response [4] for further insights.

The variety of signals and features for stress detection necessitates the use of multi-modal data fusion approaches and time-series analysis. To date, several reviews studied this topic outlining the advantages and disadvantages of various algorithms and methods for stress analysis [11], [20], [24], justifying the benefit of multi-modal approaches to estimate stress with high accuracy [11], [24]. For this reason, for the sake of completeness, we will just mention in this section some of the most common approaches that can be adopted for stress analysis, by redirecting the reader to the papers mentioned earlier for additional details. As mainly adopted algorithms we can mention: Support Vector Machines, Random Forests, Fuzzy Logic Algorithms, K-Nearest Neighbours, Logistic Regression, Naive Bayes classifiers, Ensemble Methods, Artificial Neural Networks, Convolutional Neural Networks and Recurrent Neural Networks [11], [12], [19], [20], [21], [22], [23].

The mentioned methods vary in complexity, training time, and reliance on features. However, the type and the conditions in which the data are collected before the training of any model remain crucial factors, influencing its performance. What has been said could partially justify the consistent heterogeneity in terms of performance in the literature [11], [19], [24]. All the above underlines once again the importance and need for a well-structured experimental protocol, from the assessment of the participants to data collection and data analysis. Doing so would provide more relevant data to the algorithms and, consequently, better performance in terms of stress detection.

V. CONCLUSION

One of the main limitations of a scientific study, especially when humans are involved, is the lack of reproducibility, often caused by inconsistent methodological reporting [107]. Notably, this is increasingly true when other aspects are not considered: e.g., gender imbalance in the sample, missing socio-demographic analysis and details about the conditions under which the experiment takes place. When investigating the effect of stress, additional variability is linked to the non-specific response of humans in front of stressors [2], [35], requiring further attention for the experimental protocol design.

The analysis of the ground truth is certainly the trickiest part when stress is investigated. Except for the analysis of the CORT level, which requires either salivary, hair, urinary, plasma, serum, or sweat fingernails samples [26] and can not be performed in real-time [27], defining a reliable ground truth is not possible. For this reason, labelling the acquired data as either stressful or not would implicitly enhance a bias. As reported in [4], different approaches exist to label the data, conscious of the previous limitation: the results of the questionnaires and the occurrence of a stressful situation (e.g., a cognitive task), whether considered separately or together, can be used for this purpose. However, as previously noted, the experimenter should keep in mind the limitations of each questionnaire and the possibility that the user's response may not accurately reflect a stressful condition. This once again highlights the necessity of a reliable experimental protocol.

With this paper, which combines the efforts of researchers with different backgrounds, we want to provide the reader insights on how to induce and assess stress effectively by seeing the problem from technical and psychological perspectives. Hopefully, this would bring awareness of the different aspects that should be considered when studying and investigating stress, enhancing the quality of the experiments and reducing biases.

Nevertheless, the work is not free of limitations. We focused exclusively on common cognitive tasks as approaches to induce acute mental stress. However, this approach offers the advantage of allowing researchers to quantify cognitive impairment following induced acute stress. An additional limitation is that only hardware and practical approaches were analyzed to mitigate the effects of exogenous sources, whereas software solutions were overlooked. Indeed, addressing these would require separate work.

To conclude, the next subsections will be dedicated to some final remarks related to both stress induction and assessment.

A. INDUCING STRESS

Designing a protocol that guarantees a stressful condition is fundamental, especially for all those studies that explore the effect of different treatments for stress reduction, e.g. through binaural beats [230], [231]. This is especially true if the CORT is not monitored. The great variability in subjects' stress response is a limitation in stress studies. Indeed, assuming a cognitive task to elicit enough stress in the candidates is not always correct. Note that this is true regardless of the task presented. For this reason, both personal and situational factors should be considered [49], [50], [51] to create a stressful event. The presence of an audience during data acquisition, a competition condition, rewards, punishments, evaluative conditions and no multiple attempts are approaches that the experimenter could adopt.

B. ASSESSING STRESS

The stress assessment procedure starts with an in-depth analysis of the samples under investigation. The CORT level is influenced by several factors like phase of the day, hours of sleep, sex and pharmacological treatments, that if not considered, could bias the result of an experiment. Furthermore, if cognitive tasks are leveraged to induce stress in the participants, factors like education and age must be considered as inclusion or exclusion criteria.

Despite biases, questionnaires are useful to have a rapid response from the participants about their condition. It is preferable to fill out at least two questionnaires for assessing the level of stress (or anxiety) in the investigated sample, as suggested in [58]. Having a detailed screen of the participants about the psychological condition is pivotal for a study, and will guarantee homogeneity if comparison between independent groups is required. Some questionnaires are preferable to be presented before and after the experiments, whereas others like the visual analogue scale and the self-assessment manikin can be easily presented between tasks (if the protocol includes multiple acquisitions), because of their rapidity of response.

Besides the type of signal considered for stress detection, ranging from electrical to inertial measurements, it is important to analyze carefully the conditions under which the signals are collected and the sources of inaccuracy that would limit the quality of the data. If not, the collected data will be biased, limiting the generalization of a machine or deep learning algorithm. Devices used should guarantee comfort during the experiment and suitability for long-term measurements if the next step is to test the technology outside the laboratory. It is worth noting that solutions exploiting wires and attachable electrodes may be a confounding factor, eliciting a stress response [34] during the experiment. For this reason, comfort and miniaturized devices with low incisiveness should be considered if possible.

C. FUTURE PERSPECTIVE

Emerging technologies such as Virtual Reality (VR) and Augmented Reality (AR) have a significant potential to induce acute mental stress. The realism is achieved through elements like place illusion, plausibility, and virtual body ownership [232], emphasizing VR's potential in psychological studies focused on stress induction and assessment. By designing laboratory experiments to replicate stress levels akin to those encountered in real-world settings [232], researchers may uncover previously unidentified stressors and novel psychological and physiological patterns in response to them, opening new research paths to be explored. Similarly, AR technologies can be adopted to simulate stressful situations in the real world, e.g., simulating a critical condition for training purposes [233], understanding the human response in personalized and crucial scenarios. Altogether, these technologies are set to become the new "controlled environment" for studying the effects of induced acute mental stress, most likely without the need for a physical laboratory in which to conduct the experiments.

Last but not least, particular attention should be given to chronic stress, which disrupts the immune system balance, causing peripheral and central inflammation, and contributing to various stress-related diseases [234]. Indeed, the increased cases of mental disorders in recent years require new approaches to monitor patients' emotional states, going beyond the traditional assessments typically used in a clinical environment [235]. In this context, VR, AR and smart sensors can play an important role. This is particularly true whether the combination of these systems is meant for telemedicine applications, which we believe will see widespread adoption in the coming years. The use of smart sensors like smartwatches, smart shoes and even smart mattresses [236] can be used for this purpose by monitoring the mental health of the patients toward personalized medicine, improving the quality of the treatments, or simply preventing the onset of a chronic stress condition.

REFERENCES

[1] World Health Organization. (Jun. 17, 2022). *Mental Health*. Accessed: Jul. 27, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response

[2] R. S. Lazarus and S. Folkman, *Stress, Appraisal, and Coping*. Cham, Switzerland: Springer, 1984.

[3] *The Burnout Report*, Mental Health U.K., London, U.K., Jan. 2024.

[4] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on psychological stress detection using biosignals," *IEEE Trans. Affect. Comput.*, vol. 13, no. 1, pp. 440–460, Jan. 2022.

[5] N. Schneiderman, G. Ironson, and S. D. Siegel, "Stress and health: Psychological, behavioral, and biological determinants," *Annu. Rev. Clin. Psychol.*, vol. 1, no. 1, pp. 607–628, Apr. 2005.

[6] L. Sandrini, A. Ieraci, P. Amadio, M. Zarà, and S. S. Barbieri, "Impact of acute and chronic stress on thrombosis in healthy individuals and cardiovascular disease patients," *Int. J. Mol. Sci.*, vol. 21, no. 21, p. 7818, Oct. 2020.

[7] K. L. Tamashiro, R. R. Sakai, C. A. Shively, I. N. Karatsoreos, and L. P. Reagan, "Chronic stress, metabolism, and metabolic syndrome," *Stress*, vol. 14, no. 5, pp. 468–474, Nov. 2011.

[8] G. C. Dieleman, A. C. Huizink, J. H. M. Tulen, E. M. W. J. Utens, H. E. Creemers, J. van der Ende, and F. C. Verhulst, "Alterations in HPA-axis and autonomic nervous system functioning in childhood anxiety disorders point to a chronic stress hypothesis," *Psychoneuroendocrinology*, vol. 51, pp. 135–150, Jan. 2015.

[9] AXA. (Mar. 29, 2023). *The True Cost of Running on Empty: Work-Related Stress Costing U.K. Economy 28bn a Year*. [Online]. Available: https://www.axa.co.uk/newsroom/media-releases/2023

[10] L. Whited, T. LeBouef, and Z. Yaker. (2023). *Physiology, Automic Nervous System*. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK538516

[11] S. Gedam and S. Paul, "A review on mental stress detection using wearable sensors and machine learning techniques," *IEEE Access*, vol. 9, pp. 84045–84066, 2021.

[12] M. Gil-Martin, R. San-Segundo, A. Mateos, and J. Ferreiros-Lopez, "Human stress detection with wearable sensors using convolutional neural networks," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 37, no. 1, pp. 60–70, Jan. 2022.

[13] R. Castaldo, W. Xu, P. Melillo, L. Pecchia, L. Santamaria, and C. James, "Detection of mental stress due to oral academic examination via ultra-short-term HRV analysis," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 3805–3808.

[14] S. M. Seipäjärvi, A. Tuomola, J. Juurakko, M. Rottensteiner, A.-P.-E. Rissanen, J. L. O. Kurkela, U. M. Kujala, J. A. Laukkanen, and J. Wikgren, "Measuring psychosocial stress with heart rate variability-based methods in different health and age groups," *Physiolog. Meas.*, vol. 43, no. 5, May 2022, Art. no. 055002.

[15] R. Katarya and S. Maan, "Stress detection using smartwatches with machine learning: A survey," in *Proc. Int. Conf. Electron. Sustain. Commun. Syst. (ICESC)*, Jul. 2020, pp. 306–310.

[16] A. R. Raju, R. Ramadevi, P. R. Babu, and V. D, "Galvanic skin response based stress detection system using machine learning and IoT," in *Proc. 2nd Int. Conf. Augmented Intell. Sustain. Syst. (ICAISS)*, Aug. 2023, pp. 709–714.

[17] R. F. Navea, P. J. Buenvenida, and C. D. Cruz, "Stress detection using galvanic skin response: An Android application," *J. Phys., Conf. Ser.*, vol. 1372, no. 1, Nov. 2019, Art. no. 012001.

[18] R. Martinez, A. Salazar-Ramirez, A. Arruti, E. Irigoyen, J. I. Martin, and J. Muguerza, "A self-paced relaxation response detection system based on galvanic skin response analysis," *IEEE Access*, vol. 7, pp. 43730–43741, 2019.

[19] R. K. Nath, H. Thapliyal, A. Caban-Holt, and S. P. Mohanty, "Machine learning based solutions for real-time stress monitoring," *IEEE Consum. Electron. Mag.*, vol. 9, no. 5, pp. 34–41, Sep. 2020.

[20] S. Elzeiny and M. Qaraqe, "Machine learning approaches to automatic stress detection: A review," in *Proc. IEEE/ACS 15th Int. Conf. Comput. Syst. Appl. (AICCSA)*, Oct. 2018, pp. 1–6.

[21] K. Masood and M. A. Alghamdi, "Modeling mental stress using a deep learning framework," *IEEE Access*, vol. 7, pp. 68446–68454, 2019.

[22] D. Robles, M. Benchekroun, V. Zalc, D. Istrate, and C. Taramasco, "Stress detection from surface electromyography using convolutional neural networks," in *Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2022, pp. 3235–3238.

[23] C.-Y. Liao, R.-C. Chen, and S.-K. Tai, "Emotion stress detection using EEG signal and deep learning technologies," in *Proc. IEEE Int. Conf. Appl. Syst. Invention (ICASI)*, Apr. 2018, pp. 90–93.

[24] G. Taskasaplidis, D. A. Fotiadis, and P. D. Bamidis, "Review of stress detection methods using wearable sensors," *IEEE Access*, vol. 12, pp. 38219–38246, 2024.

[25] P. Boucher and P. Plusquellec, "Acute stress assessment from excess cortisol secretion: Fundamentals and perspectives," *Frontiers Endocrinology*, vol. 10, p. 749, Nov. 2019.

[26] T. Iqbal, A. Elahi, W. Wijns, and A. Shahzad, "Cortisol detection methods for stress monitoring in connected health," *Health Sci. Rev.*, vol. 6, Mar. 2023, Art. no. 100079.

[27] M. Benchekroun, D. Istrate, V. Zalc, and D. Lenne, "MMSD: A multi-modal dataset for real-time, continuous stress detection from physiological signals," in *Proc. 15th Int. Joint Conf. Biomed. Eng. Syst. Technol.*, 2022, pp. 240–248.

[28] A. Arza, J. M. Garzón-Rey, J. Lázaro, E. Gil, R. Lopez-Anton, C. de la Camara, P. Laguna, R. Bailon, and J. Aguiló, "Measuring acute stress response through physiological signals: Towards a quantitative assessment of stress," *Med. Biol. Eng. Comput.*, vol. 57, no. 1, pp. 271–287, Jan. 2019.

[29] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, "Introducing WESAD, a multimodal dataset for wearable stress and affect detection," in *Proc. 20th ACM Int. Conf. Multimodal Interact.*, Oct. 2018, pp. 400–408.

[30] V. Markova, T. Ganchev, and K. Kalinkov, "CLAS: A database for cognitive load, affect and stress recognition," in *Proc. Int. Conf. Biomed. Innov. Appl. (BIA)*, Nov. 2019, pp. 1–4.

[31] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis ;Using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, Jan. 2012.

[32] L. Stappen, A. Baird, L. Christ, L. Schumann, B. Sertolli, E.-M. Meßner, E. Cambria, G. Zhao, and B. Schüller, "The MuSe 2021 multimodal sentiment analysis challenge: Sentiment, emotion, physiological-emotion, and stress," in *Proc. 2nd Multimodal Sentiment Anal. Challenge*, 2021, pp. 5–14.

[33] W. Chen, S. Zheng, and X. Sun, "Introducing MDPSD, a multimodal dataset for psychological stress detection," in *Proc. 8th CCF Conf.*, Chongqing, China. Cham, Switzerland: Springer, Jan. 2021, pp. 59–82.

[34] H. Chaptoukaev, V. Strizhkova, M. Panariello, B. Dalpaos, A. Reka, V. Manera, S. Thümmler, E. Ismailova, W. Nicholas, N. Evans, F. Brémond, M. Todisco, M. A. Zuluaga, and L. M. Ferrari, "StressID: A multimodal dataset for stress identification," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 36, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Eds., Red Hook, NY, USA: Curran Associates, 2023, pp. 29798–29811. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2023/file/5f09bfe6730e9627a9f800d01a8ad5cd-Paper-Datasets_and_Benchmarks.pdf

[35] H. Selye, *The Stress Life*. New York, NY, USA: McGraw-Hill, 1956.

[36] P. Karthikeyan, M. Murugappan, and S. Yaacob, "A review on stress inducement stimuli for assessing human stress using physiological signals," in *Proc. IEEE 7th Int. Colloq. Signal Process. its Appl.*, Mar. 2011, pp. 420–425.

[37] M. A. Staal, "Stress, cognition, and human performance: A literature review and conceptual framework," NASA Ames Res. Centre, MoRet Field, CA, USA, NASA Tech. Memorandum 212824, 2004.

[38] M. Egger, M. Ley, and S. Hanke, "Emotion recognition from physiological signal analysis: A review," *Electron. Notes Theor. Comput. Sci.*, vol. 343, pp. 35–55, May 2019.

[39] J. A. Russell and A. Mehrabian, "Evidence for a three-factor theory of emotions," *J. Res. Personality*, vol. 11, no. 3, pp. 273–294, Sep. 1977.

[40] J. P. Herman, J. M. McKlveen, S. Ghosal, B. L. Kopp, A. C. Wulsin, R. Makinson, J. R. Scheimann, and B. Myers, "Regulation of the hypothalamic-pituitary-adrenocortical stress response," *Comprehensive Physiol.*, vol. 6, no. 2, pp. 603–621, Mar. 2016.

[41] E. Zavala, K. C. A. Wedgwood, M. Voliotis, J. Tabak, F. Spiga, S. L. Lightman, and K. Tsaneva-Atanasova, "Mathematical modelling of endocrine systems," *Trends Endocrinol. Metabolism*, vol. 30, no. 4, pp. 244–257, Feb. 2019.

[42] A. I. Turner, N. Smyth, S. J. Hall, S. J. Torres, M. Hussein, S. U. Jayasinghe, K. Ball, and A. J. Clow, "Psychological stress reactivity and future health and disease outcomes: A systematic review of prospective evidence," *Psychoneuroendocrinology*, vol. 114, Apr. 2020, Art. no. 104599.

[43] E. Broug-Holub, J. H. A. Persoons, K. Schornagel, S. C. Mastbergen, and G. Kraal, "Effects of stress on alveolar macrophages: A role for the sympathetic nervous system," *Amer. J. Respiratory Cell Mol. Biol.*, vol. 19, no. 5, pp. 842–848, Nov. 1998.

[44] B. Sadoul and B. Geffroy, "Measuring cortisol, the major stress hormone in fishes," *J. Fish Biol.*, vol. 94, no. 4, pp. 540–555, Apr. 2019.

[45] D. Y. Lee, E. Kim, and M. H. Choi, "Technical and clinical aspects of cortisol as a biochemical marker of chronic stress," *BMB Rep.*, vol. 48, no. 4, pp. 209–216, Apr. 2015.

[46] L. Thau, J. Gandhi, and S. Sharma, "Physiology, cortisol," in *StatPearls [Internet]*. Treasure Island, FL, USA: StatPearls Publishing, Jan. 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK538239/

[47] S. Folkman and J. T. Moskowitz, "Coping: Pitfalls and promise," *Annu. Rev. Psychol.*, vol. 55, no. 1, pp. 745–774, Feb. 2004.

[48] M. Zeidner and N. S. Endler, *Handbook of Coping: Theory, Research, Applications*, vol. 195. Hoboken, NJ, USA: Wiley, 1995.

[49] R. S. Lazarus, *Stress and Emotion: A New Synthesis*. Cham, Switzerland: Springer, 2006.

[50] J. Thatcher and M. C. Day, "Re-appraising stress appraisals: The underlying properties of stress in sport," *Psychol. Sport Exercise*, vol. 9, no. 3, pp. 318–335, May 2008.

[51] R. S. Lazarus, "Emotions and interpersonal relationships: Toward a person-centered conceptualization of emotions and coping," *J. Personality*, vol. 74, no. 1, pp. 9–46, Feb. 2006.

[52] V. R. LeBlanc, "The effects of acute stress on performance: Implications for health professions education," *Academic Med.*, vol. 84, no. 10, pp. S25–S33, Oct. 2009.
[53] R. F. Baumeister and C. J. Showers, "A review of paradoxical performance effects: Choking under pressure in sports and mental tests," *Eur. J. Social Psychol.*, vol. 16, no. 4, pp. 361–383, Oct. 1986.
[54] C. D. Chapman, C. Benedict, and H. B. Schiöth, "Experimenter gender and replicability in science," *Sci. Adv.*, vol. 4, no. 1, Jan. 2018, Art. no. e1701427.
[55] A. Duchesne, E. Tessera, K. Dedovic, V. Engert, and J. C. Pruessner, "Effects of panel sex composition on the physiological stress responses to psychosocial stress in healthy young men and women," *Biol. Psychol.*, vol. 89, no. 1, pp. 99–106, Jan. 2012.
[56] B. M. Kudielka, D. H. Hellhammer, and C. Kirschbaum, "Ten years of research with the trier social stress test—Revisited," in *Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior*, E. Harmon-Jones and P. Winkielman, Eds., The Guilford Press, 2007, pp. 56–83.
[57] M. Biondi and A. Picardi, "Psychological stress and neuroendocrine function in humans: The last two decades of research," *Psychotherapy Psychosomatics*, vol. 68, no. 3, pp. 114–150, 1999.
[58] N. F. Narvaez Linares, V. Charron, A. J. Ouimet, P. R. Labelle, and H. Plamondon, "A systematic review of the trier social stress test methodology: Issues in promoting study comparison and replicable research," *Neurobiol. Stress*, vol. 13, Nov. 2020, Art. no. 100235.
[59] S. Zhong, I. Shalev, D. Koh, R. P. Ebstein, and S. H. Chew, "Competitiveness and stress," *Int. Econ. Rev.*, vol. 59, no. 3, pp. 1263–1281, 2018.
[60] 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.
[61] D. Wiswede, T. F. Münte, and J. Rüsseler, "Negative affect induced by derogatory verbal feedback modulates the neural signature of error detection," *Social Cognit. Affect. Neurosci.*, vol. 4, no. 3, pp. 227–237, Sep. 2009.
[62] S. Thuillard, M. Adams, G. Jelmini, S. Schmutz, A. Sonderegger, and J. Sauer, "When humans and computers induce social stress through negative feedback: Effects on performance and subjective state," *Comput. Hum. Behav.*, vol. 133, Aug. 2022, Art. no. 107270.
[63] R. F. Baumeister, "Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance," *J. Personality Social Psychol.*, vol. 46, no. 3, pp. 610–620, 1984.
[64] H. J. McNamara and R. I. Fisch, "Effect of high and low motivation on two aspects of attention," *Perceptual Motor Skills*, vol. 19, no. 2, pp. 571–578, Oct. 1964.
[65] L. J. Moore, S. J. Vine, M. R. Wilson, and P. Freeman, "The effect of challenge and threat states on performance: An examination of potential mechanisms," *Psychophysiology*, vol. 49, no. 10, pp. 1417–1425, Oct. 2012.
[66] W. R. Low, P. Freeman, J. Butt, M. Stoker, and I. Maynard, "The role and creation of pressure in training: Perspectives of athletes and sport psychologists," *J. Appl. Sport Psychol.*, vol. 35, no. 4, pp. 710–730, Oct. 2023.
[67] H. Ernst, M. Scherpf, S. Pannasch, J. R. Helmert, H. Malberg, and M. Schmidt, "Assessment of the human response to acute mental stress–An overview and a multimodal study," *PLoS ONE*, vol. 18, no. 11, Nov. 2023, Art. no. e0294069.
[68] P. L. Dobkin, C. Létourneau, and C. Breault, "Determining baseline and adaptation periods in stress research," *Psychotherapy Psychosomatics*, vol. 61, nos. 1–2, pp. 109–116, 1994.
[69] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Tröster, and U. Ehlert, "Discriminating stress from cognitive load using a wearable EDA device," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 410–417, Mar. 2010.
[70] R. L. Piferi, K. A. Kline, J. Younger, and K. A. Lawler, "An alternative approach for achieving cardiovascular baseline: Viewing an aquatic video," *Int. J. Psychophysiol.*, vol. 37, no. 2, pp. 207–217, Aug. 2000.
[71] S. Grassini, G. V. Segurini, and M. Koivisto, "Watching nature videos promotes physiological restoration: Evidence from the modulation of alpha waves in electroencephalography," *Frontiers Psychol.*, vol. 13, Jun. 2022, Art. no. 871143.
[72] J. Shin, J. Kwon, J. Choi, and C.-H. Im, "Ternary near-infrared spectroscopy brain–computer interface with increased information transfer rate using prefrontal hemodynamic changes during mental arithmetic, breath-holding, and idle state," *IEEE Access*, vol. 6, pp. 19491–19498, 2018.
[73] Q. Wang and O. Sourina, "Real-time mental arithmetic task recognition from EEG signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 2, pp. 225–232, Mar. 2013.
[74] P. Karthikeyan, M. Murugappan, and S. Yaacob, "A study on mental arithmetic task based human stress level classification using discrete wavelet transform," in *Proc. IEEE Conf. Sustain. Utilization Develop. Eng. Technol. (STUDENT)*, Oct. 2012, pp. 77–81.
[75] W. Linden, "What do arithmetic stress tests measure? Protocol variations and cardiovascular responses," *Psychophysiology*, vol. 28, no. 1, pp. 91–102, Jan. 1991.
[76] K. Dedovic, R. Renwick, N. K. Mahani, V. Engert, S. Lupien, and J. C. Pruessner, "The Montreal imaging stress task: Using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain," *J. Psychiatry Neurosci.*, vol. 30, no. 5, p. 319, Sep. 2005.
[77] C. W. Lejuez, C. W. Kahler, and R. A. Brown, "A modified computer version of the Paced Auditory Serial Addition Task (PASAT) as a laboratory-based stressor," *Behav. Therapist*, vol. 26, no. 4, pp. 290–293, 2003.
[78] T. Tombaugh, "A comprehensive review of the paced auditory serial addition test (PASAT)," *Arch. Clin. Neuropsychol.*, vol. 21, no. 1, pp. 53–76, Jan. 2006.
[79] T. Kolotylova, M. Koschke, K.-J. Bär, U. Ebner-Priemer, N. Kleindienst, M. Bohus, and C. Schmahl, "Development of the 'Mannheim Multicomponent Stress Test' (MMST)," *Psychotherapie, Psychosomatik, Medizinische Psychologie*, vol. 60, no. 2, pp. 64–72, 2009.
[80] T. Reinhardt, C. Schmahl, S. Wüst, and M. Bohus, "Salivary cortisol, heart rate, electrodermal activity and subjective stress responses to the Mannheim multicomponent stress test (MMST)," *Psychiatry Res.*, vol. 198, no. 1, pp. 106–111, Jun. 2012.
[81] A. Bali and A. S. Jaggi, "Clinical experimental stress studies: Methods and assessment," *Rev. Neurosciences*, vol. 26, no. 5, pp. 555–579, Oct. 2015.
[82] J. R. Stroop, "Studies of interference in serial verbal reactions," *J. Experim. Psychol.*, vol. 18, no. 6, pp. 643–662, Dec. 1935.
[83] R. Salo, A. Henik, and L. C. Robertson, "Interpreting stroop interference: An analysis of differences between task versions," *Neuropsychology*, vol. 15, no. 4, pp. 462–471, 2001.
[84] T. W. Buchanan, J. S. Laures-Gore, and M. C. Duff, "Acute stress reduces speech fluency," *Biol. Psychol.*, vol. 97, pp. 60–66, Mar. 2014.
[85] 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, nos. 1–2, pp. 76–81, 1993.
[86] J. Radun, H. Maula, V. Rajala, M. Scheinin, and V. Hongisto, "Speech is special: The stress effects of speech, noise, and silence during tasks requiring concentration," *Indoor Air*, vol. 31, no. 1, pp. 264–274, Jan. 2021.
[87] A. V. Machado, M. G. Pereira, G. G. L. Souza, M. Xavier, C. Aguiar, L. de Oliveira, and I. Mocaiber, "Association between distinct coping styles and heart rate variability changes to an acute psychosocial stress task," *Sci. Rep.*, vol. 11, no. 1, p. 24025, Dec. 2021.
[88] R. Richer, V. Koch, L. Abel, F. Hauck, M. Kurz, V. Ringgold, V. Müller, A. Küderle, L. Schindler-Gmelch, B. M. Eskofier, and N. Rohleder, "Machine learning-based detection of acute psychosocial stress from body posture and movements," *Sci. Rep.*, vol. 14, no. 1, p. 8251, Apr. 2024.
[89] A.-M. Brouwer and M. A. Hogervorst, "A new paradigm to induce mental stress: The sing-a-song stress test (SSST)," *Frontiers Neurosci.*, vol. 8, p. 224, Jul. 2014.
[90] J. T. Le, P. Watson, D. Begg, L. Albertella, and M. E. Le Pelley, "Physiological and subjective validation of a novel stress procedure: The simple singing stress procedure," *Behav. Res. Methods*, vol. 53, no. 4, pp. 1478–1487, Aug. 2021.
[91] M. M. Bradley and P. J. Lang, "The international affective picture system (IAPS) in the study of emotion and attention," in *Handbook Emotion Elicitation Assessment*, vol. 29, J. A. Coan and J. J. B. Allen, Eds., London, U.K.: Oxford Univ. Press, 2007, pp. 29–46.
[92] B. Kurdi, S. Lozano, and M. R. Banaji, "Introducing the open affective standardized image set (OASIS)," *Behav. Res. Methods*, vol. 49, no. 2, pp. 457–470, Apr. 2017.

[93] F. R. Ihmig, F. Neurohr-Parakenings, S. K. Schäfer, J. Lass-Hennemann, and T. Michael, "On-line anxiety level detection from biosignals: Machine learning based on a randomized controlled trial with spider-fearful individuals," *PLoS ONE*, vol. 15, no. 6, Jun. 2020, Art. no. e0231517.

[94] A. Buske-Kirschbaum, S. Jobst, A. Wustmans, C. Kirschbaum, W. Rauh, and D. Hellhammer, "Attenuated free cortisol response to psychosocial stress in children with atopic dermatitis," *Psychosomatic Med.*, vol. 59, no. 4, pp. 419–426, 1997.

[95] J. A. Seddon, V. J. Rodriguez, Y. Provencher, J. Raftery-Helmer, J. Hersh, P. R. Labelle, and K. Thomassin, "Meta-analysis of the effectiveness of the trier social stress test in eliciting physiological stress responses in children and adolescents," *Psychoneuroendocrinology*, vol. 116, Jun. 2020, Art. no. 104582.

[96] M. R. Gunnar, B. M. Reid, B. Donzella, Z. R. Miller, S. Gardow, N. C. Tsakonas, K. M. Thomas, M. DeJoseph, and J. J. Bendezú, "Validation of an online version of the trier social stress test in a study of adolescents," *Psychoneuroendocrinology*, vol. 125, Mar. 2021, Art. no. 105111.

[97] M. Wallergård, P. Jönsson, G. Johansson, and B. Karlson, "A virtual reality version of the trier social stress test: A pilot study," *Presence*, vol. 20, no. 4, pp. 325–336, Aug. 2011.

[98] T. Smeets, S. Cornelisse, C. W. E. M. Quaedflieg, T. Meyer, M. Jelicic, and H. Merckelbach, "Introducing the Maastricht acute stress test (MAST): A quick and non-invasive approach to elicit robust autonomic and glucocorticoid stress responses," *Psychoneuroendocrinology*, vol. 37, no. 12, pp. 1998–2008, Dec. 2012.

[99] K. Gambetta-Tessini, R. Mariño, M. Morgan, W. Evans, and V. Anderson, "Stress and health-promoting attributes in Australian, New Zealand, and Chilean dental students," *J. Dental Educ.*, vol. 77, no. 6, pp. 801–809, Jun. 2013.

[100] O. Kayikcioglu, S. Bilgin, G. Seymenoglu, and A. Deveci, "State and trait anxiety scores of patients receiving intravitreal injections," *Biomed. Hub*, vol. 2, no. 2, pp. 1–5, Aug. 2017.

[101] F. Dutheil, B. Pereira, F. Moustafa, G. Naughton, F.-X. Lesage, and C. Lambert, "At-risk and intervention thresholds of occupational stress using a visual analogue scale," *PLoS ONE*, vol. 12, no. 6, Jun. 2017, Art. no. e0178948.

[102] G. Balasubramanian, A. Kanagasabai, J. Mohan, and N. P. G. Seshadri, "Music induced emotion using wavelet packet decomposition—An EEG study," *Biomed. Signal Process. Control*, vol. 42, pp. 115–128, Apr. 2018.

[103] A. D. Prabaswari, C. Basumerda, and B. W. Utomo, "The mental workload analysis of staff in study program of private educational organization," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 528, no. 1, May 2019, Art. no. 012018.

[104] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: The PANAS scales," *J. Personality Social Psychol.*, vol. 54, no. 6, pp. 1063–1070, 1988.

[105] E. R. Thompson, "Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS)," *J. Cross-Cultural Psychol.*, vol. 38, no. 2, pp. 227–242, Mar. 2007.

[106] S. Sinha and G. S. Latha, "Coping response to same stressors varies with gender," *Nat. J. Physiol. Pharm Pharmacol.*, vol. 7, pp. 1053–1057, Jan. 2018.

[107] C. Tannenbaum, R. P. Ellis, F. Eyssel, J. Zou, and L. Schiebinger, "Sex and gender analysis improves science and engineering," *Nature*, vol. 575, no. 7781, pp. 137–146, Nov. 2019.

[108] C. Kirschbaum, B. M. Kudielka, J. Gaab, N. C. Schommer, and D. H. Hellhammer, "Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis," *Psychosomatic Med.*, vol. 61, no. 2, pp. 154–162, 1999.

[109] K. L. Felmingham, W. C. Fong, and R. A. Bryant, "The impact of progesterone on memory consolidation of threatening images in women," *Psychoneuroendocrinology*, vol. 37, no. 11, pp. 1896–1900, Nov. 2012.

[110] D. Sharp, M. Cole, C. Lave, H. P. Ginsburg, A. L. Brown, and L. A. French, "Education and cognitive development: The evidence from experimental research," *Monographs Soc. Res. Child Develop.*, vol. 44, no. 1/2, p. 1, 1979.

[111] A. C. Incollingo Rodriguez, E. S. Epel, M. L. White, E. C. Standen, J. R. Seckl, and A. J. Tomiyama, "Hypothalamic-pituitary-adrenal axis dysregulation and cortisol activity in obesity: A systematic review," *Psychoneuroendocrinology*, vol. 62, pp. 301–318, Dec. 2015.

[112] J. Fine, K. L. Branan, A. J. Rodriguez, T. Boonya-Ananta, Ajmal, J. C. Ramella-Roman, M. J. McShane, and G. L. Coté, "Sources of inaccuracy in photoplethysmography for continuous cardiovascular monitoring," *Biosensors*, vol. 11, no. 4, p. 126, Apr. 2021.

[113] L. Mesin, "Volume conductor models in surface electromyography: Computational techniques," *Comput. Biol. Med.*, vol. 43, no. 7, pp. 942–952, Aug. 2013.

[114] T. M. Burke, R. R. Markwald, A. W. McHill, E. D. Chinoy, J. A. Snider, S. C. Bessman, C. M. Jung, J. S. O'Neill, and K. P. Wright, "Effects of caffeine on the human circadian clock in vivo and in vitro," *Sci. Transl. Med.*, vol. 7, no. 305, Sep. 2015, Art. no. 305ra146.

[115] W. Dimpfel, F. Schober, and M. Spüler, "The influence of caffeine on human EEG under resting condition and during mental loads," *Clin. Investigator*, vol. 71, no. 3, pp. 197–207, Mar. 1993.

[116] U. M. Nater, N. Rohleder, W. Schlotz, U. Ehlert, and C. Kirschbaum, "Determinants of the diurnal course of salivary alpha-amylase," *Psychoneuroendocrinology*, vol. 32, no. 4, pp. 392–401, May 2007.

[117] D. B. Boivin and P. Boudreau, "Impacts of shift work on sleep and circadian rhythms," *Pathologie Biologie*, vol. 62, no. 5, pp. 292–301, Oct. 2014.

[118] M. Balbo, R. Leproult, and E. Van Cauter, "Impact of sleep and its disturbances on hypothalamo-pituitary-adrenal axis activity," *Int. J. Endocrinol.*, vol. 2010, pp. 1–16, Jan. 2010.

[119] H. G. Lund, B. D. Reider, A. B. Whiting, and J. R. Prichard, "Sleep patterns and predictors of disturbed sleep in a large population of college students," *J. Adolescent Health*, vol. 46, no. 2, pp. 124–132, Feb. 2010.

[120] J.-P. Chaput, C. Dutil, and H. Sampasa-Kanyinga, "Sleeping hours: What is the ideal number and how does age impact this?" *Nature Sci. Sleep*, vol. 10, pp. 421–430, Nov. 2018.

[121] C. J. Patrick, J. J. Curtin, and A. Tellegen, "Development and validation of a brief form of the multidimensional personality Questionnaire," *Psychol. Assessment*, vol. 14, no. 2, pp. 150–163, 2002.

[122] E. Childs, T. L. White, and H. de Wit, "Personality traits modulate emotional and physiological responses to stress," *Behavioural Pharmacol.*, vol. 25, no. 5, pp. 493–502, 2014.

[123] C.-A. Chun, R. H. Moos, and R. C. Cronkite, "Culture: A fundamental context for the stress and coping paradigm," in *Handbook of Multicultural Perspectives on Stress and Coping*, P. T. P. Wong and L. C. J. Wong, Eds., Spring Publications, 2006, pp. 29–53, doi: 10.1007/0-387-26238-5_2.

[124] H. Lee, T. Masuda, K. Ishii, Y. Yasuda, and Y. Ohtsubo, "Cultural differences in the perception of daily stress between European Canadian and Japanese undergraduate students," *Personality Social Psychol. Bull.*, vol. 49, no. 4, pp. 571–584, Apr. 2023.

[125] R. Miller and C. Kirschbaum, "Cultures under stress: A cross-national meta-analysis of cortisol responses to the trier social stress test and their association with anxiety-related value orientations and internalizing mental disorders," *Psychoneuroendocrinology*, vol. 105, pp. 147–154, Jul. 2019.

[126] B. Van der Kolk, "Posttraumatic stress disorder and the nature of trauma," *Dialogues Clin. Neurosci.*, vol. 2, no. 1, pp. 7–22, 2000.

[127] C. Demetriou, B. U. Ozer, and C. A. Essau, "Self-report questionnaires," in *The Encyclopedia of Clinical Psychology*, R. Cautin and S. Lilienfeld, Eds., Hoboken, NJ, USA: Wiley, 2015, pp. 1–6.

[128] S. Cohen, "Perceived stress in a probability sample of the United States," in *The Social Psychology of Health*, S. Spacapan and S. Oskamp, Eds., Newbury Park, CA, USA: Sage, 1988, pp. 31–67.

[129] C. Spielberger, R. Gorsuch, R. Lushene, P. R. Vagg, and G. Jacobs, *Manual for the State-Trait Anxiety Inventory (Form Y1–Y2)*. Palo Alto, CA, USA: Consulting Psychologists Press.

[130] L. Klimek, K.-C. Bergmann, T. Biedermann, J. Bousquet, P. Hellings, K. Jung, H. Merk, H. Olze, W. Schlenter, P. Stock, J. Ring, M. Wagenmann, W. Wehrmann, R. Mösges, and O. Pfaar, "Visual analogue scales (VAS): Measuring instruments for the documentation of symptoms and therapy monitoring in cases of allergic rhinitis in everyday health care: Position paper of the German society of allergology (AeDA) and the German society of allergy and clinical immunology (DGAKI), ENT section, in collaboration with the working group on clinical immunology, allergology and environmental medicine of the German society of otorhinolaryngology, head and neck surgery (DGHNOKHC)," *Allergo J. Int.*, vol. 26, no. 1, pp. 16–24, Feb. 2017.

[131] A. M. Mitchell, P. A. Crane, and Y. Kim, "Perceived stress in survivors of suicide: Psychometric properties of the perceived stress scale," *Res. Nursing Health*, vol. 31, no. 6, pp. 576–585, Dec. 2008.

[132] F.-X. Lesage, S. Berjot, and F. Deschamps, "Clinical stress assessment using a visual analogue scale," *Occupational Med.*, vol. 62, no. 8, pp. 600–605, Dec. 2012.
[133] A. J. Roth, A. B. Kornblith, L. Batel-Copel, E. Peabody, H. I. Scher, and J. C. Holland, "Rapid screening for psychologic distress in men with prostate carcinoma: A pilot study," *Cancer*, vol. 82, no. 10, pp. 1904–1908, May 1998.
[134] I. L. D. Houtman and F. C. Bakker, "The anxiety thermometer: A validation study," *J. Personality Assessment*, vol. 53, no. 3, pp. 575–582, Sep. 1989.
[135] O. Thomas, S. Hanton, and G. M. Jones, "An alternative approach to short-form self-report assessment of competitive anxiety: A research note," *Int. J. Sport Psychol.*, vol. 33, no. 3, pp. 325–336, Jan. 2002.
[136] V. Tran, "Positive affect negative affect scale (PANAS)," in *Encyclopedia of Behavioral Medicine*. Cham, Switzerland: Springer, 2020, pp. 1708–1709.
[137] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *J. Behav. Therapy Experim. Psychiatry*, vol. 25, no. 1, pp. 49–59, Mar. 1994.
[138] S. G. Hart, "Nasa-task load index (NASA-TLX); 20 years later," in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, Los Angeles, CA, USA, Jan. 2006, pp. 904–908.
[139] M. R. Wilson, J. M. Poolton, N. Malhotra, K. Ngo, E. Bright, and R. S. W. Masters, "Development and validation of a surgical workload measure: The surgery task load index (SURG-TLX)," *World J. Surg.*, vol. 35, no. 9, pp. 1961–1969, Sep. 2011.
[140] D. Harris, M. Wilson, and S. Vine, "Development and validation of a simulation workload measure: The simulation task load index (SIM-TLX)," *Virtual Reality*, vol. 24, no. 4, pp. 557–566, Dec. 2020.
[141] C. A. Latkin, C. Edwards, M. A. Davey-Rothwell, and K. E. Tobin, "The relationship between social desirability bias and self-reports of health, substance use, and social network factors among urban substance users in Baltimore, Maryland," *Addictive Behaviors*, vol. 73, pp. 133–136, Oct. 2017.
[142] Y. Cho, N. Bianchi-Berthouze, and S. J. Julier, "DeepBreath: Deep learning of breathing patterns for automatic stress recognition using low-cost thermal imaging in unconstrained settings," in *Proc. 7th Int. Conf. Affect. Comput. Intell. Interact. (ACII)*, Oct. 2017, pp. 456–463.
[143] M. Gärtner, S. Grimm, and M. Bajbouj, "Frontal midline theta oscillations during mental arithmetic: Effects of stress," *Frontiers Behav. Neurosci.*, vol. 9, Apr. 2015, Art. no. 133588.
[144] D. Schoofs, D. Preuß, and O. T. Wolf, "Psychosocial stress induces working memory impairments in an n-back paradigm," *Psychoneuroendocrinology*, vol. 33, no. 5, pp. 643–653, Jun. 2008.
[145] M. Qi, H. Gao, and G. Liu, "The effect of mild acute psychological stress on attention processing: An ERP study," *Exp. Brain Res.*, vol. 236, no. 7, pp. 2061–2071, Jul. 2018.
[146] M. Qi, H. Gao, L. Guan, G. Liu, and J. Yang, "Subjective stress, salivary cortisol, and electrophysiological responses to psychological stress," *Frontiers Psychol.*, vol. 7, p. 229, Feb. 2016.
[147] R. Bruyer and M. Brysbaert, "Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)?" *Psychologica Belgica*, vol. 51, no. 1, p. 5, Feb. 2011.
[148] L. Mesin, G. E. Cipriani, and M. Amanzio, "Electroencephalography-based brain–machine interfaces in older adults: A literature review," *Bioengineering*, vol. 10, no. 4, p. 395, Mar. 2023.
[149] S. Chatterjee, R. S. Thakur, R. N. Yadav, L. Gupta, and D. K. Raghuvanshi, "Review of noise removal techniques in ECG signals," *IET Signal Process.*, vol. 14, no. 9, pp. 569–590, Dec. 2020.
[150] K. T. Sweeney, T. E. Ward, and S. F. McLoone, "Artifact removal in physiological signals—Practices and possibilities," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 488–500, May 2012.
[151] E. Huigen, A. Peper, and C. A. Grimbergen, "Investigation into the origin of the noise of surface electrodes," *Med. Biol. Eng. Comput.*, vol. 40, no. 3, pp. 332–338, May 2002.
[152] B. Bent, B. A. Goldstein, W. A. Kibbe, and J. P. Dunn, "Investigating sources of inaccuracy in wearable optical heart rate sensors," *npj Digit. Med.*, vol. 3, no. 1, p. 18, Feb. 2020.
[153] M. D. Peláez-Coca, A. Hernando, J. Lázaro, and E. Gil, "Impact of the PPG sampling rate in the pulse rate variability indices evaluating several fiducial points in different pulse waveforms," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 2, pp. 539–549, Feb. 2022.
[154] M. E. Dawson, A. M. Schell, and D. L. Filion, "The electrodermal system," in *Handbook of Psychophysiology*, J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson, Eds., 3rd ed., Cambridge, U.K.: Cambridge Univ. Press, 2007, pp. 159–181, doi: 10.1017/CBO9780511546396.007.
[155] W. Boucsein, D. C. Fowles, S. Grimnes, G. Ben-Shakhar, W. T. Roth, M. E. Dawson, and D. L. Filion, "Publication recommendations for electrodermal measurements," *Psychophysiology*, vol. 49, no. 8, pp. 1017–1034, Aug. 2012.
[156] H. F. Posada-Quintero, J. P. Florian, A. D. Orjuela-Cañón, T. Aljama-Corrales, S. Charleston-Villalobos, and K. H. Chon, "Power spectral density analysis of electrodermal activity for sympathetic function assessment," *Ann. Biomed. Eng.*, vol. 44, no. 10, pp. 3124–3135, Oct. 2016.
[157] Y. Shimomura, T. Yoda, K. Sugiura, A. Horiguchi, K. Iwanaga, and T. Katsuura, "Use of frequency domain analysis of skin conductance for evaluation of mental workload," *J. Physiolog. Anthropol.*, vol. 27, no. 4, pp. 173–177, 2008.
[158] R. Merletti and G. L. Cerone, "Tutorial. Surface EMG detection, conditioning and pre-processing: Best practices," *J. Electromyogr. Kinesiol.*, vol. 54, Oct. 2020, Art. no. 102440.
[159] I. Campanini, A. Merlo, C. Disselhorst-Klug, L. Mesin, S. Muceli, and R. Merletti, "Fundamental concepts of bipolar and high-density surface EMG understanding and teaching for clinical, occupational, and sport applications: Origin, detection, and main errors," *Sensors*, vol. 22, no. 11, p. 4150, May 2022.
[160] J. A. Urigüen and B. García-Zapirain, "EEG artifact removal—State-of-the-art and guidelines," *J. neural Eng.*, vol. 12, no. 3, Apr. 2015, Art. no. 031001.
[161] Y. Zhao and J. H. M. Bergmann, "Non-contact infrared thermometers and thermal scanners for human body temperature monitoring: A systematic review," *Sensors*, vol. 23, no. 17, p. 7439, Aug. 2023.
[162] V. Shusterman, K. P. Anderson, and O. Barnea, "Spontaneous skin temperature oscillations in normal human subjects," *Amer. J. Physiol.-Regulatory, Integrative Comparative Physiol.*, vol. 273, no. 3, pp. R1173–R1181, Sep. 1997.
[163] W. Karlen, S. Raman, J. M. Ansermino, and G. A. Dumont, "Multiparameter respiratory rate estimation from the photoplethysmogram," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 7, pp. 1946–1953, Jul. 2013.
[164] S. Sarkar, S. Bhattacherjee, and S. Pal, "Extraction of respiration signal from ECG for respiratory rate estimation," in *Proc. Michael Faraday IET Int. Summit*, Sep. 2015, pp. 336–340.
[165] D. J. Meredith, D. Clifton, P. Charlton, J. Brooks, C. W. Pugh, and L. Tarassenko, "Photoplethysmographic derivation of respiratory rate: A review of relevant physiology," *J. Med. Eng. Technol.*, vol. 36, no. 1, pp. 1–7, Mar. 2012.
[166] C. Massaroni, A. Nicolò, D. Lo Presti, M. Sacchetti, S. Silvestri, and E. Schena, "Contact-based methods for measuring respiratory rate," *Sensors*, vol. 19, no. 4, p. 908, Feb. 2019.
[167] D. Jarchi, D. Salvi, L. Tarassenko, and D. A. Clifton, "Validation of instantaneous respiratory rate using reflectance PPG from different body positions," *Sensors*, vol. 18, no. 11, p. 3705, Oct. 2018.
[168] S. Mathôt and A. Vilotijević, "Methods in cognitive pupillometry: Design, preprocessing, and statistical analysis," *Behav. Res. Methods*, vol. 55, no. 6, pp. 3055–3077, Aug. 2022.
[169] C. Kelbsch, T. Strasser, Y. Chen, B. Feigl, P. D. Gamlin, R. Kardon, T. Peters, K. A. Roecklein, S. R. Steinhauer, E. Szabadi, A. J. Zele, H. Wilhelm, and B. J. Wilhelm, "Standards in pupillography," *Frontiers Neurol.*, vol. 10, p. 129, Feb. 2019.
[170] S. R. Steinhauer, M. M. Bradley, G. J. Siegle, K. A. Roecklein, and A. Dix, "Publication guidelines and recommendations for pupillary measurement in psychophysiological studies," *Psychophysiology*, vol. 59, no. 4, pp. e14035, Apr. 2022.
[171] D. Byrne et al., "An international comparison of long-term average speech spectra," *J. Acoust. Soc. Amer.*, vol. 96, no. 4, pp. 2108–2120, Oct. 1994.
[172] V. Best, S. Carlile, C. Jin, and A. van Schaik, "The role of high frequencies in speech localization," *J. Acoust. Soc. Amer.*, vol. 118, no. 1, pp. 353–363, Jul. 2005.
[173] F. Busquet, F. Efthymiou, and C. Hildebrand, "Voice analytics in the wild: Validity and predictive accuracy of common audio-recording devices," *Behav. Res. Methods*, vol. 56, no. 3, pp. 2114–2134, May 2023.

[174] F. Bo, M. Yerebakan, Y. Dai, W. Wang, J. Li, B. Hu, and S. Gao, "IMU-based monitoring for assistive diagnosis and management of IoHT: A review," *Healthcare*, vol. 10, no. 7, p. 1210, Jun. 2022.

[175] R. Khusainov, D. Azzi, I. Achumba, and S. Bersch, "Real-time human ambulation, activity, and physiological monitoring: Taxonomy of issues, techniques, applications, challenges and limitations," *Sensors*, vol. 13, no. 10, pp. 12852–12902, Sep. 2013.

[176] B. Altinöz and D. Ünsal, "Determining efficient temperature test points for IMU calibration," in *Proc. IEEE/ION Position, Location Navigat. Symp. (PLANS)*, Apr. 2018, pp. 552–556.

[177] D. Tedaldi, A. Pretto, and E. Menegatti, "A robust and easy to implement method for IMU calibration without external equipments," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2014, pp. 3042–3049.

[178] P. Kligfield, L. S. Gettes, J. J. Bailey, R. Childers, B. J. Deal, E. W. Hancock, G. Van Herpen, J. A. Kors, P. Macfarlane, and D. M. Mirvis, "Recommendations for the standardization and interpretation of the electrocardiogram: Part I: The electrocardiogram and its technology: A scientific statement from the American Heart Association electrocardiography and arrhythmias committee, council on clinical cardiology; the American college of cardiology foundation; and the heart rhythm society endorsed by the international society for computerized electrocardiology," *Circulation*, vol. 115, no. 10, pp. 1306–1324, 2007.

[179] A. Kamal, J. Harness, G. Irving, and A. Mearns, "Skin photoplethysmography—A review," *Comput. Methods Programs Biomed.*, vol. 28, no. 4, pp. 257–269, 1989.

[180] O. Kwon, J. Jeong, H. B. Kim, I. H. Kwon, S. Y. Park, J. E. Kim, and Y. Choi, "Electrocardiogram sampling frequency range acceptable for heart rate variability analysis," *Healthcare Informat. Res.*, vol. 24, no. 3, p. 198, 2018.

[181] F. Esgalhado, A. Batista, V. Vassilenko, S. Russo, and M. Ortigueira, "Peak detection and HRV feature evaluation on ECG and PPG signals," *Symmetry*, vol. 14, no. 6, p. 1139, Jun. 2022.

[182] E. Mejía-Mejía, J. M. May, M. Elgendi, and P. A. Kyriacou, "Differential effects of the blood pressure state on pulse rate variability and heart rate variability in critically ill patients," *NPJ Digit. Med.*, vol. 4, no. 1, p. 82, May 2021.

[183] W.-H. Lin, D. Wu, C. Li, H. Zhang, and Y. Zhang, "Comparison of heart rate variability from PPG with that from ECG," in *Proc. Int. Conf. Health Informat.*, Vilamoura, Portugal. Cham, Switzerland: Springer, Nov. 2013, pp. 213–215.

[184] N. Pinheiro, R. Couceiro, J. Henriques, J. Muehlsteff, I. Quintal, L. Gonçalves, and P. Carvalho, "Can PPG be used for HRV analysis?" in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 2945–2949.

[185] E. Mejía-Mejía, J. M. May, R. Torres, and P. A. Kyriacou, "Pulse rate variability in cardiovascular health: A review on its applications and relationship with heart rate variability," *Physiolog. Meas.*, vol. 41, no. 7, Aug. 2020, Art. no. 07TR01.

[186] E. Gil, M. Orini, R. Bailón, J. M. Vergara, L. Mainardi, and P. Laguna, "Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions," *Physiolog. Meas.*, vol. 31, no. 9, pp. 1271–1290, Sep. 2010.

[187] A. Guala, C. Camporeale, L. Ridolfi, and L. Mesin, "Non-invasive aortic systolic pressure and pulse wave velocity estimation in a primary care setting: An in silico study," *Med. Eng. Phys.*, vol. 42, pp. 91–98, Apr. 2017.

[188] D. A. Tong, K. A. Bartels, and K. S. Honeyager, "Adaptive reduction of motion artifact in the electrocardiogram," in *Proc. 2nd Joint 24th Annu. Conf. Annu. Fall Meeting Biomed. Eng. Soc. Eng. Med. Biol.*, vol. 2, 2002, pp. 1403–1404.

[189] D. Berwal, S. Dewan, and M. S. Baghini, "Motion artifact removal in ambulatory ECG signal for heart rate variability analysis," *IEEE Sensors J.*, vol. 19, no. 24, pp. 12432–12442, Dec. 2019.

[190] M. R. Ram, K. V. Madhav, E. H. Krishna, N. R. Komalla, and K. A. Reddy, "A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 5, pp. 1445–1457, May 2012.

[191] A. Baranchuk, J. Kang, C. Shaw, D. Campbell, S. Ribas, W. M. Hopman, H. Alanazi, D. P. Redfearn, and C. S. Simpson, "Electromagnetic interference of communication devices on ECG machines," *Clin. Cardiol.*, vol. 32, no. 10, pp. 588–592, Oct. 2009.

[192] V. Hartmann, H. Liu, F. Chen, Q. Qiu, S. Hughes, and D. Zheng, "Quantitative comparison of photoplethysmographic waveform characteristics: Effect of measurement site," *Frontiers Physiol.*, vol. 10, p. 198, Mar. 2019.

[193] T. Röddiger, C. Clarke, P. Breitling, T. Schneegans, H. Zhao, H. Gellersen, and M. Beigl, "Excerpt from 'sensing with earables: A systematic literature review and taxonomy of phenomena,"' in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, Sep. 2022, vol. 6, no. 3, pp. 244–245.

[194] E.-S. Väliaho, J. A. Lipponen, P. Kuoppa, T. J. Martikainen, H. Jäntti, T. T. Rissanen, M. Castrén, J. Halonen, M. P. Tarvainen, T. M. Laitinen, T. P. Laitinen, O. E. Santala, O. Rantula, N. S. Naukkarinen, and J. E. K. Hartikainen, "Continuous 24-h photoplethysmogram monitoring enables detection of atrial fibrillation," *Frontiers Physiol.*, vol. 12, Jan. 2022, Art. no. 778775.

[195] E. Nemati, M. J. Deen, and T. Mondal, "A wireless wearable ECG sensor for long-term applications," *IEEE Commun. Mag.*, vol. 50, no. 1, pp. 36–43, Jan. 2012.

[196] Y.-D. Lee and W.-Y. Chung, "Wireless sensor network based wearable smart shirt for ubiquitous health and activity monitoring," *Sens. Actuators B, Chem.*, vol. 140, no. 2, pp. 390–395, Jul. 2009.

[197] V. G. Sirtoli, M. Liamini, L. T. Lins, M. Lessard-Tremblay, G. E. R. Cowan, R. J. Zednik, and G. Gagnon, "Removal of motion artifacts in capacitive electrocardiogram acquisition: A review," *IEEE Trans. Biomed. Circuits Syst.*, vol. 17, no. 3, pp. 394–412, Jun. 2023.

[198] D. Naranjo, R. Cattaneo, and L. Mesin, "Development of a prototype for the analysis of multiple responses of the autonomic nervous system," *Biomed. Signal Process. Control*, vol. 70, Sep. 2021, Art. no. 102994.

[199] C. M. Laine, K. M. Spitler, C. P. Mosher, and K. M. Gothard, "Behavioral triggers of skin conductance responses and their neural correlates in the primate amygdala," *J. Neurophysiol.*, vol. 101, no. 4, pp. 1749–1754, Apr. 2009.

[200] H. F. Posada-Quintero, J. P. Florian, Á. D. Orjuela-Cañón, and K. H. Chon, "Highly sensitive index of sympathetic activity based on time-frequency spectral analysis of electrodermal activity," *Amer. J. Physiol.-Regulatory, Integrative Comparative Physiol.*, vol. 311, no. 3, pp. R582–R591, Sep. 2016.

[201] M.-B. Hossain, H. F. Posada-Quintero, Y. Kong, R. McNaboe, and K. H. Chon, "Automatic motion artifact detection in electrodermal activity data using machine learning," *Biomed. Signal Process. Control*, vol. 74, Apr. 2022, Art. no. 103483.

[202] M. S. Qasim, D. S. Bari, and Ø. G. Martinsen, "Influence of ambient temperature on tonic and phasic electrodermal activity components," *Physiolog. Meas.*, vol. 43, no. 6, Jun. 2022, Art. no. 065001.

[203] M. van Dooren, J. J. G. J. de Vries, and J. H. Janssen, "Emotional sweating across the body: Comparing 16 different skin conductance measurement locations," *Physiol. Behav.*, vol. 106, no. 2, pp. 298–304, May 2012.

[204] M.-Z. Poh, N. C. Swenson, and R. W. Picard, "A wearable sensor for unobtrusive, long-term assessment of electrodermal activity," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 5, pp. 1243–1252, May 2010.

[205] D. Kamińska, K. Smółka, and G. Zwoliński, "Detection of mental stress through EEG signal in virtual reality environment," *Electronics*, vol. 10, no. 22, p. 2840, Nov. 2021.

[206] E. Perez-Valero, M. A. Vaquero-Blasco, M. A. Lopez-Gordo, and C. Morillas, "Quantitative assessment of stress through EEG during a virtual reality stress-relax session," *Frontiers Comput. Neurosci.*, vol. 15, Jul. 2021, Art. no. 684423.

[207] J. Wijsman, B. Grundlehner, J. Penders, and H. Hermens, "Trapezius muscle EMG as predictor of mental stress," *ACM Trans. Embedded Comput. Syst.*, vol. 12, no. 4, pp. 1–20, Jun. 2013.

[208] S. Orguc, H. S. Khurana, K. M. Stankovic, H. S. Leel, and A. P. Chandrakasan, "EMG-based real time facial gesture recognition for stress monitoring," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2018, pp. 2651–2654.

[209] A. B. Usakli, "Improvement of EEG signal acquisition: An electrical aspect for state of the art of front end," *Comput. Intell. Neurosci.*, vol. 2010, pp. 1–7, Jan. 2010.

[210] J. M. DeFreitas, T. W. Beck, and M. S. Stock, "Comparison of methods for removing electromagnetic noise from electromyographic signals," *Physiolog. Meas.*, vol. 33, no. 2, pp. 147–158, Feb. 2012.

[211] D. Seok, S. Lee, M. Kim, J. Cho, and C. Kim, "Motion artifact removal techniques for wearable EEG and PPG sensor systems," *Frontiers Electron.*, vol. 2, May 2021, Art. no. 685513.

[212] C. H. Vinkers, R. Penning, J. Hellhammer, J. C. Verster, J. H. G. M. Klaessens, B. Olivier, and C. J. Kalkman, "The effect of stress on core and peripheral body temperature in humans," *Stress*, vol. 16, no. 5, pp. 520–530, Sep. 2013.

[213] M. Gjoreski, M. Luštrek, M. Gams, and H. Gjoreski, "Monitoring stress with a wrist device using context," *J. Biomed. Informat.*, vol. 73, pp. 159–170, Sep. 2017.

[214] Y. S. Can, N. Chalabianloo, D. Ekiz, J. Fernandez-Alvarez, G. Riva, and C. Ersoy, "Personal stress-level clustering and decision-level smoothing to enhance the performance of ambulatory stress detection with smartwatches," *IEEE Access*, vol. 8, pp. 38146–38163, 2020.

[215] K. A. Herborn, J. L. Graves, P. Jerem, N. P. Evans, R. Nager, D. J. McCafferty, and D. E. F. McKeegan, "Skin temperature reveals the intensity of acute stress," *Physiol. Behav.*, vol. 152, pp. 225–230, Dec. 2015.

[216] D. I. Sessler, A. Moayeri, R. Støen, B. Glosten, J. Hynson, and J. McGuire, "Thermoregulatory vasoconstriction decreases cutaneous heat loss," *Anesthesiology*, vol. 73, no. 4, pp. 656–660, 1990.

[217] I. Fernández-Cuevas, J. C. B. Marins, J. A. Lastras, P. M. G. Carmona, S. P. Cano, M. Á. García-Concepción, and M. Sillero-Quintana, "Classification of factors influencing the use of infrared thermography in humans: A review," *Infr. Phys. Technol.*, vol. 71, pp. 28–55, Jul. 2015.

[218] L. Mesin, A. Monaco, and R. Cattaneo, "Investigation of nonlinear pupil dynamics by recurrence quantification analysis," *BioMed Res. Int.*, vol. 2013, no. 1, pp. 1–11, 2013.

[219] F. Ponzio, A. E. L. Villalobos, L. Mesin, C. de'Sperati, and S. Roatta, "A human–computer interface based on the 'voluntar' pupil accommodative response," *Int. J. Hum.-Comput. Stud.*, vol. 126, pp. 53–63, Jun. 2019.

[220] P. Ren, A. Barreto, J. Huang, Y. Gao, F. R. Ortega, and M. Adjouadi, "Off-line and on-line stress detection through processing of the pupil diameter signal," *Ann. Biomed. Eng.*, vol. 42, no. 1, pp. 162–176, Jan. 2014.

[221] M. Pedrotti, M. A. Mirzaei, A. Tedesco, J.-R. Chardonnet, F. Mérienne, S. Benedetto, and T. Baccino, "Automatic stress classification with pupil diameter analysis," *Int. J. Hum.-Comput. Interact.*, vol. 30, no. 3, pp. 220–236, Mar. 2014.

[222] A. Pomè, D. C. Burr, A. Capuozzo, and P. Binda, "Spontaneous pupillary oscillations increase during mindfulness meditation," *Current Biol.*, vol. 30, no. 18, pp. R1030–R1031, Sep. 2020.

[223] B. Winn, D. Whitaker, D. B. Elliott, and N. J. Phillips, "Factors affecting light-adapted pupil size in normal human subjects," *Investigative Ophthalmol. Vis. Sci.*, vol. 35, no. 3, p. 11327, Mar. 1994.

[224] H. Kurniawan, A. V. Maslov, and M. Pechenizkiy, "Stress detection from speech and galvanic skin response signals," in *Proc. 26th IEEE Int. Symp. Comput.-Based Med. Syst.*, Jun. 2013, pp. 209–214.

[225] K. Tomba, J. Dumoulin, E. Mugellini, O. Abou Khaled, and S. Hawila, "Stress detection through speech analysis," in *Proc. 15th Int. Joint Conf. E-Business Telecommun.*, 2018, pp. 560–564.

[226] M. Zubair, C. Yoon, H. Kim, J. Kim, and J. Kim, "Smart wearable band for stress detection," in *Proc. 5th Int. Conf. IT Converg. Secur. (ICITCS)*, Aug. 2015, pp. 1–4.

[227] T. B. Tang, L. W. Yeo, and D. J. H. Lau, "Activity awareness can improve continuous stress detection in galvanic skin response," in *Proc. IEEE SENSORS*, Nov. 2014, pp. 1980–1983.

[228] F.-T. Sun, C. Kuo, H.-T. Cheng, S. Buthpitiya, P. Collins, and M. Griss, "Activity-aware mental stress detection using physiological sensors," in *Proc. 2nd Int. ICST Conf. Mobile Comput., Appl., Services*, Santa Clara, CA, USA. Cham, Switzerland: Springer, Jan. 2012, pp. 211–230.

[229] E. Garcia-Ceja, V. Osmani, and O. Mayora, "Automatic stress detection in working environments from smartphones' accelerometer data: A first step," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 4, pp. 1053–1060, Jul. 2016.

[230] Y. Badr, F. Al-Shargie, U. Tariq, F. Babiloni, F. Al-Mughairbi, and H. Al-Nashash, "Mental stress detection and mitigation using machine learning and binaural beat stimulation," in *Proc. 45th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2023, pp. 1–5.

[231] R. Katmah, F. Al-Shargie, U. Tariq, F. Babiloni, F. Al-Mughairbi, and H. Al-Nashash, "Mental stress management using fNIRS directed connectivity and audio stimulation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1086–1096, 2023.

[232] I. Ladakis, D. Filos, and I. Chouvarda, "Virtual reality environments for stress reduction and management: A scoping review," *Virtual Reality*, vol. 28, no. 1, p. 50, Mar. 2024.

[233] J. Stuart, I. Akinnola, F. Guido-Sanz, M. Anderson, D. Diaz, G. Welch, and B. Lok, "Applying stress management techniques in augmented reality: Stress induction and reduction in healthcare providers during virtual triage simulation," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces Abstr. Workshops (VRW)*, Mar. 2020, pp. 171–172.

[234] Y.-Z. Liu, Y.-X. Wang, and C.-L. Jiang, "Inflammation: The common pathway of stress-related diseases," *Frontiers Hum. Neurosci.*, vol. 11, Jun. 2017, Art. no. 273283.

[235] P. Bufano, M. Laurino, S. Said, A. Tognetti, and D. Menicucci, "Digital phenotyping for monitoring mental disorders: Systematic review," *J. Med. Internet Res.*, vol. 25, Dec. 2023, Art. no. e46778.

[236] N. Carbonaro, M. Laurino, A. Greco, C. Marinai, F. Giannetti, F. Righetti, F. D. Rienzo, G. Rho, L. Arcarisi, M. Zanoletti, P. Bufano, M. Tesconi, N. Sgambelluri, D. Menicucci, C. Vallati, and A. Tognetti, "Smart sensors for daily-life data collection toward precision and personalized medicine: The TOLIFE project approach," in *Proc. Medit. Conf. Med. Biol. Eng. Comput.*, Carlotta Marinai, France. Cham, Switzerland: Springer, Jan. 2024, pp. 783–794.

MATTEO RAGGI received the B.Sc. and M.Sc. degrees in biomedical engineering from the Politecnico di Torino, in 2019 and 2022, respectively, where he is currently pursuing the Ph.D. degree in the electrical, electronics and communications engineering with the Department of Electronics and Telecommunications. He is also a member of the Mathematical Biology and Physiology Group (Department of Electronics and Telecommunications). His main research interests include biomedical signal processing and wearable devices.

LEE J. MOORE received the B.Sc. degree (Hons) in exercise and sports science, the M.Sc. degree in sport and health sciences, and the Ph.D. degree in examining the effects of psychophysiological responses to stress (i.e., challenge and threat states) on motor performance from the University of Exeter. He is currently a Chartered Psychologist with the British Psychological Society and a Senior Lecturer in sport and performance psychology with the University of Bath. He is an experimental Psychologist, interested in performance variability under stress, and how stress influences cognitive appraisals, emotional and physiological responses, visual attention, and motor performance.

LUCA MESIN received the degree in electronics engineering and the Ph.D. degree in applied mathematics from the Politecnico di Torino, Italy, in 1999 and 2003, respectively. He is currently an Associate Professor in biomedical engineering and a Supervisor of the Mathematical Biology and Physiology Group, Department of Electronics and Telecommunications, Politecnico di Torino. His current research interests include biomedical image and signal processing and mathematical modeling.