IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 31, 2019, accepted November 20, 2019, date of publication December 3, 2019, date of current version December 17, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2957357

New Tools and Methods in Selection of Air Traffic Controllers Based on Multimodal Psychophysiological Measurements

KRE IMIR OSI [1], SINI A POPOVI [1], (Member, IEEE), MARKO ARLIJA [1], IGOR MIJI [1], MIRKO KOKOT [1], IVAN KESED I [1], GARY STRANGMAN [2], VLADIMIR IVKOVI [2], AND QUAN ZHANG [2], (Member, IEEE)

[1]Laboratory for Interactive Simulation Systems, Faculty of Electrical Engineering and Computing, University of Zagreb, 10000 Zagreb, Croatia
[2]Neural Systems Group, Department of Psychiatry, Harvard Medical School and Massachusetts General Hospital, Charlestown, MA 02129, USA

Corresponding author: Siniša Popović (sinisa.popovic@fer.hr)

This work was supported by the NATO's Emerging Security Challenges Division in the Framework of the Science for Peace and Security Programme.

ABSTRACT Comprehensive multimodal psychophysiological measurements and smart data analysis based on wearable and low-cost technologies could enhance traditional air traffic controller (ATC) selection process. Many recent studies in neuro-cognitive science and stress resilience illustrated effectiveness of these multimodal measurements and appropriate metrics in comprehensive assessment of ATCs' mental states, such as cognitive workload, cognitive decline, attention deficit, fatigue, emotional and behavioural problems, etc. Accordingly, this article is focused on innovation efforts in ATC selection protocols based on a set of comprehensive stimuli and corresponding multimodal psychophysiological measurements. The concept of enhancement of ATC selection process presented in this article includes complex physiological, oculometric and speech measurements and appropriate metrics. From these multimodal measurements during specific stimulation tasks, which include different versions of acoustic startle stimuli, airblasts, semantically relevant aversive images and sounds, different versions of Stroop tests, visual tracking test, a complex set of multimodal-multidimensional features is computed as predictors of ATC candidates' future performance, like: stress resilience, workload capacity, attention, visual performance, working memory etc. Such cost-effective, more objective, non-invasive preliminary measurements, lasting no longer than 45 minutes may have good discriminative power and might be used in ATC selection processes as enhancement of current selection procedures. Comprehensive analysis of presented multimodal features during different experimental conditions might also be very useful in selection processes of other stressful professional jobs, like first responders, pilots, astronauts etc.

INDEX TERMS ATC selection, multimodal physiology, performance assessment, resilience assessment, cognitive capacity, attention deficit, fatigue, speech and oculometric features.

I. INTRODUCTION

During their operational duties, air traffic controllers (ATCs) must be capable to perform multiple stressful tasks simultaneously, must have good situational awareness which enables early perception and prediction of potential aircraft collisions with tragic consequences, must make right decisions under such conditions in a split of second and solve tough problems in stressful situations. It means that ATC professionals should be selected among candidates who are able to stay calm under the complex cognitive stress and make right decisions while processing different types of sensory information. Therefore, their profession nowadays is regarded as one of the most stressful occupations, which also requires outstanding skills and knowledge. ATC selection process is looking for individuals who can retain high level of attention, focus, cognitive capabilities, situational awareness and emotional stability under the pressure because potential human

The associate editor coordinating the review of this manuscript and approving it for publication was Vincenzo Piuri.

errors in such situations might cause catastrophic failures and tragic consequences. Only a few percent of individuals in the general population possess the required combination of these traits. This means that ATC professional job is for individuals with such extensive skills and, therefore, their selection is extremely interdisciplinary and multidisciplinary research topic. Therefore, the core of our concept of multimodal psychophysiological measurements in ATC selection is search for individuals who are capable to retain high cognitive capacities and emotional stability during high arousal emotional excitation within stressful working environment. These abilities need to be assessed during their selection and training process to get individuals who are capable to cope with all aspects of their future professional job. According to Eurocontrol, ATCs in a recent survey said that they are unafraid of their work, not nervous about problems they encounter and able to relax when they need to [1]. Their rigorous selection process and training enable them to handle all the problems which they may face, and they are always supported by their colleagues who are ready to help them.

ATC candidate performance assessment during the selection process should include the following criteria: good stress resistance, i.e. abilities to successfully handle stressful air traffic situations; high cognitive capacities; emotional stability; multi-tasking capabilities, i.e. simultaneous processing of multiple events; efficient real-time reasoning and correct decision making; good visual perception; good selective, divided and sustained attention; controllable alertness in conflicting multi-tasking situations; good spatial orientation and working memory; good mental arithmetic and logical reasoning; high physical fitness; vigilance; ability to maintain high cognitive processing capabilities after 6 or 8 hours working hours; high motivation; differentiation of relevant from irrelevant events; good planning ability and task sorting according to safety priorities; robust circadian rhythm and high metabolic resources; fluency in English; team-working skills within a complex and uncertain environment [2]–[5]. ATC candidates must also undergo a comprehensive medical examination to be medically eligible for such a tough job. The selection process of such resilient individuals who have the mental toughness to withstand the highly stressful job should minimise drop-out rate during their professional career, later on. But, traditional ATC selection processes focused on a variety of self-reports, questionnaires and interviews, provided by experienced psychologists and ATC instructors, have well-known limitations related to the intrinsic biases and subjective human nature. On the contrary, neurophysiological measurements of the human performance envelope might detect risky drifts towards safety boundary conditions and identify invalid dynamic and functional responses [6]. Ideally, through simultaneous psychological screening and real-time multimodal measurements based on comprehensive set of physiological, oculometric and speech features presented in the article we can better estimate if individual examinee is telling true or faking, or is exhausted with some complex cognitive tasks, or is stressed etc. The proposed augmentation of selection process with our comprehensive relatively short multimodal psychophysiological measurements can significantly improve and enhance traditional selection processes and protocols. Therefore, our research has been focused on evaluation of discriminative and predictive power of different stimuli and multimodal features in selection of the best ATC candidates for further training and education to get individuals with highest professional standards. Having high-performance employees, less absenteeism and less employee turnover are important goals and objectives, which could bring new quality and efficiency into organisational ATC selection process, increasing safety and saving time and money.

Proposed cost-effective and non-invasive multimodal measurements and features based on physiological, speech and oculometric dynamic features have been analysed in this article during a variety of stress-inducing stimuli and cognitive tasks. Design of relevant stimulation paradigms and selection of specific multimodal output features which provide reliable predictive information regarding candidates' training and on-the-job future performance is an important scientific challenge. Multimodal psychophysiological response features regularly include: heart rate reaction to stress, heart rate recovery after stress, heart rate variability, respiratory sinus arrhythmia, root mean square of the successive heart beat differences, phasic and tonic components of skin conductance, electromyogram and electrodermal activity based acoustic startle response, like startle reactivity and startle habituation and discrimination of startle responses in danger vs. safety experimental conditions, eye blinks, saccades, fixations, pupil dilation and constriction, voice fundamental frequency, energy, jitter, shimmer, formants, zero-crossing rate, cepstral coefficients etc. Various manifestations of stress, cognitive load and fatigue on human biological signals, physiological measures like heart rate variability, visual attention, saccadic dynamics or pupillary responses is well researched topic [7]–[9]. Oculometric analyses, i.e. eye movements, blinking patterns, and pupil diameter changes under various visual and cognitive tasks may also offer valuable insight into ATC candidates' mental capabilities [10]–[12]. A variety of speech features, like prosodic features, fundamental frequency, energy etc. computed during different candidates' mental states and conditions, like stress, cognitive load or fatigue, might be used as reliable predictors or indicators of ATC candidates' performance during their professional life cycle. These types of speech features affected by phenomena like stress, fatigue, or cognitive load might be used not only during selection processes but also during normal or shift duty operating hours as real-time indicators of their fatigue or cognitive overload. Such speech features can be used in real operational environment as prevention tools which can minimise risks of catastrophic decision-making failures due to ATC operator's cognitive overload, fatigue, sleepiness etc. For example, speech/voice data features have emerged as a non-invasive, objective measure of stress, fatigue or cognitive overload [13]–[21] and should be extremely valuable in applications in which spoken communication is already extensively used. It should also be noted that speech features generally remain the most viable approach due to its low-cost and accessibility amongst the all modalities. Generally, it is also well known that resilient individuals may have a shorter startle habituation period than vulnerable ones and may be less prone to effects of fatigue [22], [23].

Selection of optimised experimental paradigms which elicit corresponding multimodal features that reflect the impact of stress or cognitive overload on sympathetic or parasympathetic activities, autonomic nervous system balance or autonomic regulatory capacity, is a prerequisite for efficient multimodal experimentation in ATC selection process.

The input stimulation tasks and paradigms are usually related to: well-established generic stressful emotional stimuli, like acoustic startle, airblast, semantically relevant aversive images and sounds, fear-potentiated startle, prepulse inhibition etc.; a variety of cognitive tasks with different workload intensity, like multiple versions of Stroop test [24] which induce different levels of cognitive load and divided attention, e.g. basic Stroop test with different variations that are more attuned with Virtual Reality Stroop Test [25], or the Emotional Stroop test [26]; a variety of serious games and training simulators, etc.

The observed multimodal-multidimensional features variability and their statistical discriminative power justify proposed research efforts toward improvements and enhancement of traditional ATC selection process with these methods. Such comprehensive multimodal stimuli and multimodal metrics may have significant relevance especially in applications in which stress or cognitive overload may have huge impact on human performance, like among ATCs. This concept in the prediction of individual performance under the heavy stress should also be very attractive for many other highly stressful professional occupations, like pilots, first responders, astronauts, military personnel, etc. Finding the dominant set of stimulation tasks and multimodal features, that are the most relevant for differentiating resilient ATC candidates from vulnerable individuals, is a main objective of our research, so far [27]–[34].

Finally, many research studies demonstrate the applicability and the effectiveness of such neurophysiological measurements in assessing human mental states such as mental workload, cognition, emotion, fatigue, attention, vigilance, cognitive self-control [6], [33], [35]–[41]. As a consequence, proposed multimodal neurophysiological measurements may play a key role in future research on the performance of ATC operators in their working environments. Analyses of neurophysiological signals and corresponding features have the potential of providing reliable information about ATC operators' mental states, for example, if the operator's workload is exceeding his/her cognitive capacity, or if some kind of incapacitation is occurring, and can be used as valuable predictive information about ATC candidate's abilities to cope with such challenges in stressful professional life, in combination with First European Air Traffic Controller Selection Test (FEAST) and Dynamic Air Traffic Controller Radar Test (DART), developed and managed by Eurocontrol (http://feast-info.eurocontrol.int/).

II. METHODS

Design and development of new tools and methods for ATC candidates' performance assessment during the selection process should reflect the full spectrum of realistic operational demands by relevant stimuli and tasks. Low-cost wearable micro-sensors for measurements of the individual's multimodal physiological, oculomotor, and speech reactions [29], [42] have the potential to be used as an augmentation of traditional selection of ATC candidates. Accordingly, the laboratory version of our multimodal system for their performance assessment during the selection process enables measurements and analysis of multimodal responses to specific stressful stimuli and tasks, related to physiological, oculometric, and speech features [28], [30]–[33], [44], [43] (Fig. 1).

This experimental research had been approved by the Croatian Air Traffic Control authorities and the Ethical Committee of the University of Zagreb Faculty of Electrical Engineering and Computing. Signed informed consent was obtained from each individual participating in this research. The Croatian Air Traffic Control selected forty individuals for participation in our research protocol from the pool of more than one thousand applicants based on educational level, age, cognitive, perceptual and physical abilities, personality traits, vocational interests, psychological interviews, performance on simulated exercises etc. These ATC candidates (35 male, 5 female, mean age 23.97, standard deviation 2.12) underwent the three-segment multimodal stimulation protocol related to:

- **Stress Resilience:** a 15-minute stimulation paradigm for assessment of stress resilience while the candidate's peripheral physiology was continuously recorded, including ECG, EDA, respiration, and eyeblink EMG signals;
- **Visual Attention and Fatigue:** a 10-minute stimulation paradigm for assessment of visual attention and fatigue via specific visual tasks, while the candidate's eye movements were continuously recorded;
- **Cognitive Load:** a 7-minute stimulation paradigm for induction of different levels of cognitive load via established cognitive interference tasks that require verbal responses, while the candidate's voice signal was continuously recorded.

Stimulation paradigms and results related to each of the three segments are described in the following sections. Prior to exposure of the ATC candidates to all administered stimuli and tasks, they were familiarised with the details of the experimental protocol.

III. STRESS RESILIENCE ASSESSMENT BASED ON PHYSIOLOGICAL FEATURES

According to the extensive and growing research literature, stress resilience represents a complex

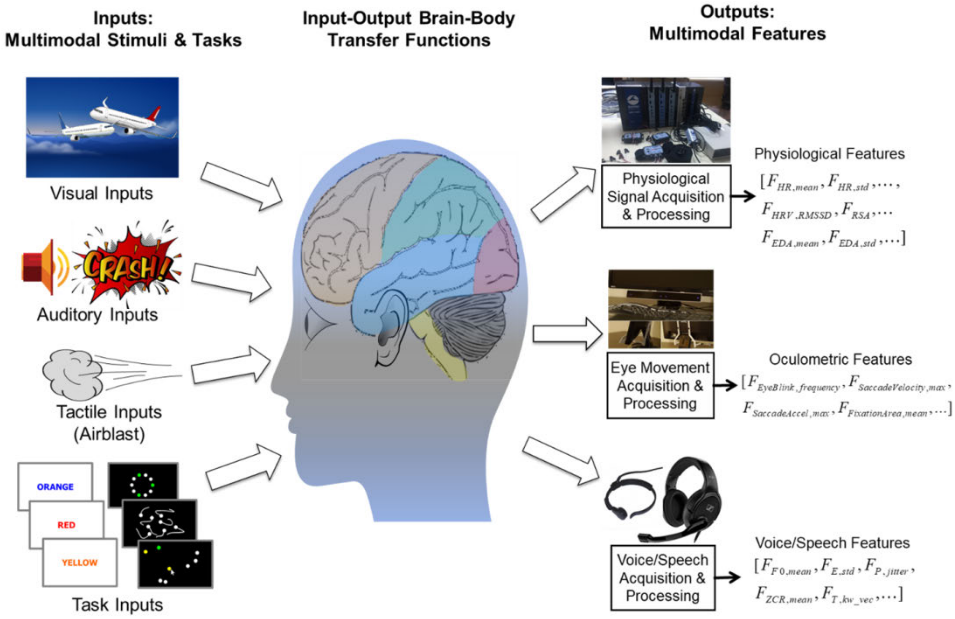

FIGURE 1. **The laboratory version of the input-output multimodal system for selection performance assessment IOMS-SPA, adapted from [44] (presented illustration was partially assembled from public domain/free sources: https://publicdomainvectors.org, http://www.stockunlimited.com).**

multimodal-multidimensional and multidisciplinary biological, cognitive, emotional and behavioural phenomenon which should be assessed by our cluster of multimodal features. While it is well-recognized that resilience definitions mostly include concepts of adversity and positive adaptation [45], it can be more generally defined as an ability to maintain normal psychological, physiological and physical functioning when exposed to extraordinary levels of stress and trauma [46]. Specifically, definitions that link resilience to the maintenance of normal functioning are particularly important to stress resilience prediction in the context of many different stressful professions. Such definition broadly implies that stress resilience could be regarded as an important contributor to operational performance under stress, as well as to healthy, long and prosperous career in such professional occupations. Generally compatible with such functional definitions of resilience is the notion that relates resilience to an ability to avoid accumulation of allostatic load following exposure to stress over some limited time horizon, where allostatic load can be measured via specific multidisciplinary features in a variety of processes in the human organism that are homeostatically regulated, relative to normal feature oscillations over time due to circadian rhythms and seasonal changes [47]. Compatibility of stress resilience definitions from allostatic-load and functional standpoint is based on findings that allostatic load contributes to impairments in cognitive performance as well as overall health over longer term [48], and might affect performance even in short-term stressful military training settings [49]. The obtained results in the previous literature illustrate discriminative power and predictive potential of several physiological features to separate resilient individuals from vulnerable ones [22], [23]. Furthermore, in our most recent research [44], the following physiological features for objectivization of stress resilience assessment confirmed discriminative power between resilient and control participants:

- **Respiratory sinus arrhythmia (RSA)**, which measures heart rate variability (HRV) in phase with inhalation and exhalation [50]. Higher measure of RSA should indicate higher resilience to stress.
- **Startle reactivity (SR)**, which measures the strength of reflexive defensive responding to an aversive unconditioned stimulus, i.e., abrupt, loud noise [51]. Lower startle reactivity should indicate higher resilience to stress.
- **Cardiac allostasis (CA)**, which measures adaptive reaction to a stressful event, involving a vigorous cardiac response to stress coupled with a significant cardiac recovery in the aftermath [52]. Higher cardiac allostasis should indicate higher resilience to stress.

A. STIMULI PARADIGMS AND TASKS

The stimulation paradigm for stress resilience assessment, lasting 15 minutes, has been designed to elicit appropriate physiological responses that are needed for computation of resilience-related features and included the following blocks:

- **Resting block.** This block lasts 3 minutes and is primarily used for computation of resting cardiac activity features, in this work RSA.
- **Block of auditory startle (AS) stimuli.** This 5-minute block is used for the elicitation and measurement of various AS-based features. In this work, it is particularly important for the assessment of general startle

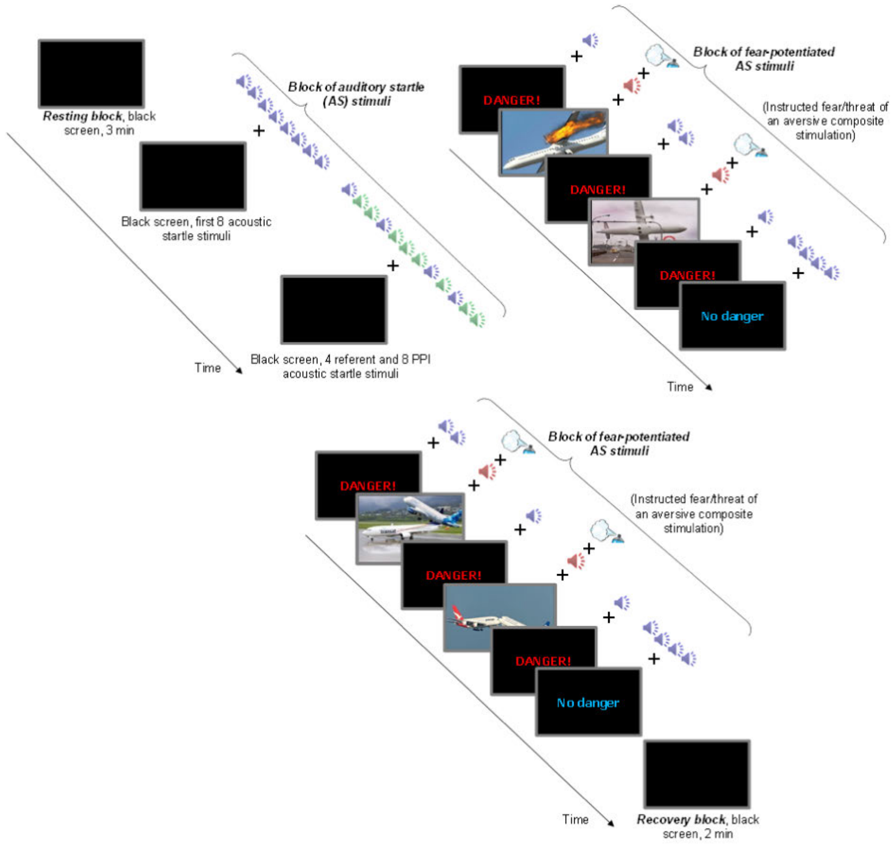

FIGURE 2. **A schematic overview of the stimulation paradigm used in resilience assessment of ATC candidates, adapted from [44].**

reactivity (SR), but various AS-based features that are related to habituation and prepulse inhibition (PPI) of the AS can be analysed as well [44].

- **Block of fear-potentiated AS stimuli.** This 5-minute block is divided into 4 phases, where the candidates are alternately confronted either with words "danger", written in red, or words for "no danger", written in blue, on a screen in front of them, and is important for the assessment of the fear-potentiated startle (FPS) response and danger vs. safety discrimination (DSD). However, the threat of composite stimuli (combinations of aversive 250-ms airblasts, loud unpleasant sounds and aversive pictures [31]) in the "danger" phases cause a strong cardiac and EDA response [44]. Hence, ECG data from this block (specifically the "danger" phases) is used for computation of the "cardiac reaction" part in the assessment of the CA feature.
- **Recovery block.** This block lasts 2 minutes and is used for the assessment of cardiac recovery from laboratory stress (the "danger" phases of the previous block). ECG data from this block is used for computation of the "cardiac recovery" part in the assessment of the CA feature.

Further details regarding the specific phases of the described paradigm, such as stimulus delivery and measurement equipment descriptions, would exceed the scope of this paper, and are described in [44]. Fig. 2 shows the timeline of the stimulation paradigm for stress resilience assessment that was used in our experiment. AS stimuli and sound components of the composite stimuli were delivered binaurally through headphones (Sennheiser PC 360 G4ME). Airblasts were directed at the back of the neck. Visual components of the composite stimuli and "danger"/"no danger" instructions were presented on a screen that was black otherwise. Just before subject's enrolment in the experiment, an experimenter verbally explained the study procedures to each participant.

B. MEASUREMENTS AND RESULTS

Biopac MP150 system with all accompanying modality-specific modules was used for collecting the ECG, EMG, EDA and respiratory data, at a sampling frequency of 1000 Hz. Gazepoint GP3 HD eye-tracker was used for spontaneous blink detection, collecting data at a frequency of 150 Hz. The synchronisation of the stimulus delivery and data acquisition hardware was performed by our IOMS-SPA software. A more detailed description of the hardware of our laboratory system can be found in [31].

After pre-processing of physiological signals (inter-beat interval (IBI) time-series based on the detected QRS complexes, filtered respiratory and EMG data), 3 selected features were computed for each ATC candidate in the following way:

- RSA – Four complementary RSA features were calculated from the accordingly processed respiratory and ECG data collected during the Resting block (3 minutes at the beginning of the paradigm), including: mean peak-to-trough difference in IBI times per respiratory cycle; time-domain correlation between the baseline-corrected IBI time series and respiratory signal; relative HRV spectral power around the breathing frequency; frequency-domain correlation between HRV and breathing. These methods have been validated and explained in detail in [44]. Final RSA estimate for each candidate is the PCA-based fusion of the four individual RSA features.
- SR – The first 4 normalised EMG-based AS responses (Block of auditory startle (AS) stimuli) have been used for the assessment of SR [44]. The average amplitude of the EMG responses preceding the spontaneous blinks during the Resting block is used as a reference.
- CA – We measured cardiac reaction by the reduction in mean IBI time, and HRV, from baseline to "danger" phases (Block of fear-potentiated AS stimuli), and cardiac recovery by the increase in mean IBI and HRV, from "danger" to recovery phase (Recovery block). Final CA estimate for each candidate is the PCA-based fusion of the four individual cardiac reaction and cardiac recovery features.

In Fig. 3, we present the two most distinct ATC candidates concerning each of the three computed features, illustrating the inter-candidate variability in the three selected features.

The discriminative power of physiological RSA, SR and CA features, illustrated by inter-candidate variability in Fig. 3, enables selection of the most resilient individuals from a group of ATC candidates. Therefore, it is reasonable to assume that the subset of these 3 features, which were statistically significant discriminators between resilient and control group [44], might have applicable predictive power in ATC selection process.

IV. VISUAL ATTENTION AND FATIGUE ASSESSMENT BASED ON OCULOMETRIC FEATURES

Healthy oculomotor system is vitally important for the success of ATCs in their professional job that require maintenance of high oculomotor capabilities under stress, cognitive overload and fatigue. Among several different types of eye movements [53], [54], a few types particularly important for calculation of relevant oculometric performance features for ATCs include saccadic and smooth pursuit movements. Saccadic eye movements abruptly orient the eyes on the target located in the periphery of the visual field. Velocity of a saccade cannot be controlled voluntarily and dominantly depends on the distance of the target from the centre of the visual field [53]; however, saccade velocity can be impacted by the direction of saccade [55] and can be slowed down by fatigue, drugs, or different pathological states [53]. The smooth-pursuit movements enable the eyes to track smoothly moving target, i.e. to keep such target in the centre of the visual field. Beside these types of eye movements, important quality of healthy oculomotor system is an ability to maintain the stationary target of interest in the centre of the visual field, which requires an active fixation system that inhibits eye movements when stationary target is examined [53].

From the described eye movements a series of oculometric features can be calculated using appropriate eye tracking technology, like: saccadic peak velocity and acceleration, saccadic trajectory deviation, saccadic latency, fixation duration, fixation area etc. Saccadic peak velocity has been particularly shown to change with varying levels of stress and fatigue [56]. It was proposed to be an index of arousal based on a review of relevant experimental evidence [55], which showed that saccadic velocities can be decreased by lower arousal and increased by higher arousal. Analyses of more complex eye movements related to spatio-temporal sequences of saccades, fixations and smooth-pursuit movements allow the study of the subject's screen attention and scan order, which has been used in research on situational awareness, information-seeking and decision-making behaviour in stressful occupations, like ATCs and pilots [57]–[61].

In addition to the eye movements themselves, in occupations requiring high cognitive performance under stress, like air traffic control, it is also important to observe the number of blinks per minute and the change of the pupil diameter. The number of blinks per minute is associated with learning processes and at least partly modulated by the levels of dopamine in the brain [62], [63]. Spontaneous blinks occur due to the physiological need to maintain moisture of the eyeball, but their rate of occurrence and duration can be affected by a variety of factors, for example [64]: visually demanding tasks can decrease blink rate and higher sympathetic nervous system activity increases it; blinks of longer duration can indicate alertness degradation and blinks of shorter duration are associated with increased visual load; higher blink rate and long eyelid closure were shown to be predictors of error in fatigued pilots. Pupil diameter is controlled by noradrenergic neurons from the locus coeruleus nucleus [65], [66], which is a major node that promotes noradrenergic signaling in the brain during stress response [67] and is involved in regulation of arousal and autonomic activity [68]. Furthermore, blink analyses and pupillometry have been used in air traffic control human factors research, particularly focusing on workload assessment [35], [40].

Based on the presented literature on the relevance of eye tracking for air traffic control research, we were focused on measurement of oculometric features calculated from the dynamics of gaze patterns, i.e. saccades and fixations, as well as specific visual attention performance measures, in response to relatively generic and demanding visual tasks

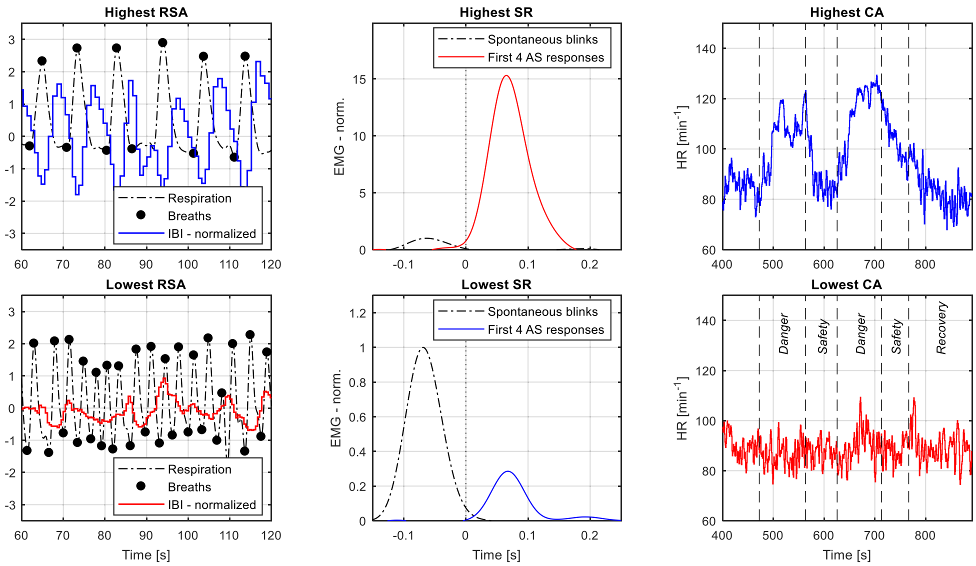

FIGURE 3. **Visualisation of the inter-candidate variability by the three selected features: RSA (left), SR (middle) and CA (right). Relevant data for the highest-scoring (top) and lowest-scoring (bottom) candidate are illustrated. Adapted from [44].**

of short duration. Our goal was to test the visual attention abilities of candidates and to generate fatigue with stimulation paradigm lasting up to 10 minutes, which would be efficient enough for application in selection processes of ATCs.

A. STIMULI PARADIGMS AND TASKS

Eye tracking stimulation paradigm applied to ATC candidates consisted of two types of tasks, alternately arranged in seven segments. Each segment of the first type of task, called Reflex Dot Pursuit (RDP), lasted for 1 minute and was used for the assessment of saccade- and fixation-related features. This task was motivated by prior research works that assessed random saccadic eye movements [10], [69]. During RDP task, the candidate was instructed to track the red dot as it consecutively appeared across the display at random locations, whenever the eye tracking system registered that the candidate's gaze reached the dot.

Segments of the second type of task, called Multiple Object Tracking (MOT), lasted for 2 minutes each and were related to single and multiple object tracking at three difficulty levels. This task was introduced based on: understanding that ''ATC domain-relevant abilities'' include breadth of visual attention [4]; solid literature foundation in the area of multiple object tracking [11], [70]–[72]; and, particularly, evidence that air-traffic control experts demonstrated better multiple object tracking performance than novices [73]. MOT task consisted of tracking 1 or 3 initially highlighted dots from the total of 8 or 10 dots during their continuous motion, which included mutual collisions.

Altogether, the experimental paradigm lasted 10 minutes and consisted of the following segments:

1. RDP task (1 minute);
2. MOT task, difficulty level I (2 minutes) – 8 subtasks related to tracking 1 dot from the total of 8 dots;

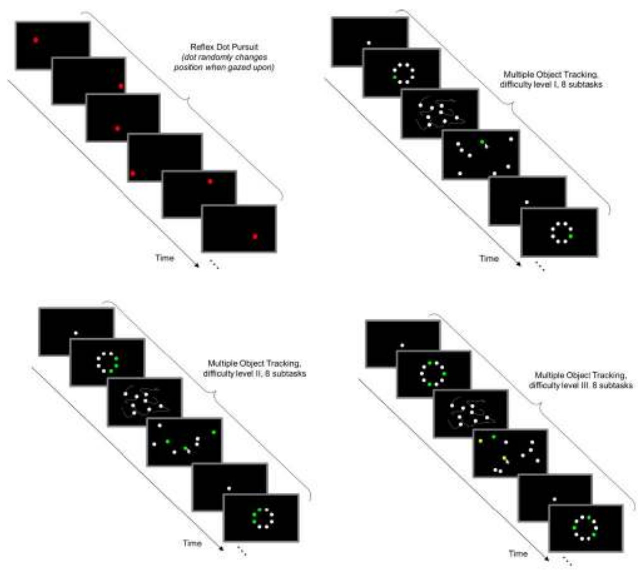

FIGURE 4. **RDP and MOT visualisation.**

3. RDP task (1 minute);
4. MOT task, difficulty level II (2 minutes) – 8 subtasks related to tracking 3 adjacent dots from a total of 8 dots;
5. RDP task (1 minute);
6. MOT task, difficulty level III (2 minutes) – 8 subtasks related to tracking 3 non-adjacent dots from a total of 10 dots;
7. RDP task (1 minute).

Illustrations of RDP and MOT tasks at different difficulty levels are presented in Fig. 4.

B. MEASUREMENTS AND RESULTS

Our research efforts related to the oculometric features and eye gaze dynamic analyses of forty ATC candidates in

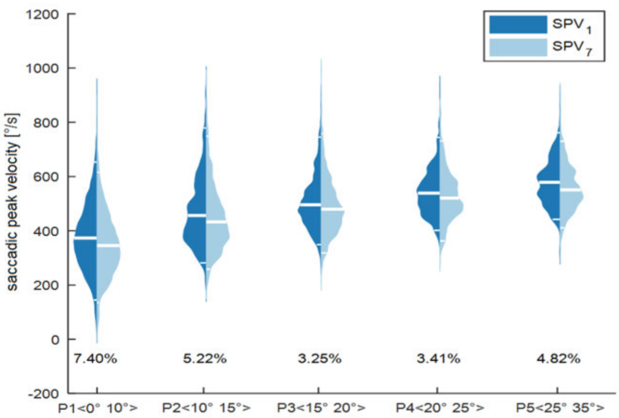

FIGURE 5. Saccadic peak velocities during the first and the last RDP segment among ATC candidates, sorted into bins depending on the visual angle.

the selection process has been done with the specialised eye tracking equipment Gazepoint GP3 HD and stimulation paradigms related to the described RDP and MOT tasks. Data acquisition frequency was 150 Hz.

Data collected during the first and last segment of RDP were compared to see if there exist measurable changes in oculometric features. In each 1-minute RDP segment, participants managed to fixate at over 200 red dots. Initial data processing consisted of detecting and classifying fixations, saccades, and valid reflex pursuits. Valid reflex pursuits were determined by gaze latency, which in the case of reflexive saccades is between 80 and 250 ms [74]. From pre-processed data, the following oculometric features were calculated for each valid reflex pursuit: saccadic peak velocity (SPV), i.e. maximum velocity calculated as the distance between two subsequent readings of saccade movement multiplied with the rate of data acquisition; saccadic peak acceleration (SPA), i.e. maximum acceleration calculated as the difference of subsequent saccadic velocities multiplied with the rate of data acquisition; saccadic trajectory deviation (STD), i.e. measured trajectory length divided by Euclidian distance between centres of fixations; fixation area (FA), i.e. the area bounded by convex hull of all gaze readings within the perimeter of the gazed dot; fixation duration (FD), i.e duration of subsequent gaze readings within the perimeter of the gazed dot.

When plotting saccade feature visualisations, saccades were sorted into bins depending on the visual angle between subsequent fixations, since it is known that saccadic velocity cannot be controlled voluntarily, but instead depends on the visual angle [53]. The following visual angle bins were applied: P1⟨0° 10°⟩, P2⟨10° 15°⟩, P3⟨15° 20°⟩, P4⟨20° 25°⟩, P5⟨25° 35°⟩.

Fig. 5 shows distributions of SPV for ATC candidates. On each violin plot, 25^{th}-50^{th}-75^{th} percentiles are marked with horizontal white lines. It can be generally seen that distribution of SPV_7, i.e. saccadic peak velocity during the last, 7^{th}, segment of the paradigm, is a few percent lower compared to saccadic peak velocity distribution from the beginning of the experiment, i.e. during the 1^{st} segment (SPV_1). At the bottom of each violin plot, it is shown how much the SPV is reduced from the beginning to the end of experiment per bin, expressed as a percentage. Similar results were obtained for saccadic peak acceleration, which is expected since it is a derivative of SPV, while the remaining three features did not show any notable differences between the first and the last RDP segment.

MOT results for each of the 3 difficulty levels were scored separately while keeping track of correct and incorrect answers. Distribution of correct answers for ATC candidates in Fig. 6 shows that ATC candidates' accuracy decreases and variance of the scores increases with the higher task difficulty.

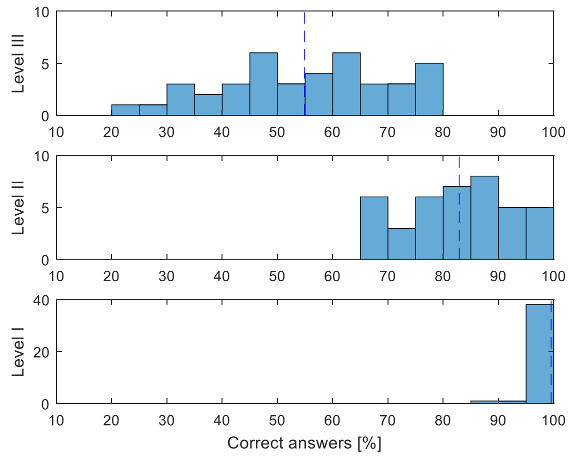

FIGURE 6. Histogram of ATC candidates' correct answers for the 3 difficulty levels of MOT.

During both RDP and MOT, a change in pupil diameter was measured in order to index stress experienced by the candidates. Based on results presented in Table 1, accuracy of ATC candidates on MOT 'difficulty level II' segment was negatively correlated with the average pupil diameter during MOT 'difficulty level II' segment (PD_4) normalised by the average pupil diameter before MOT segments, i.e. during the initial RDP segment of the stimulation paradigm (PD_1) ($p < 0.01$). Analogous results were obtained for performance of ATC candidates on the MOT 'difficulty level III' segment and normalised average pupil diameter during MOT 'difficulty

TABLE 1. Correlations between the number of correct answers on specific MOT segments and in total (i.e. performance indices) and the corresponding normalised average pupil diameters (i.e. stress indices), among ATC candidates.

Data 1	Data 2	ATC candidates	
		r	p
Difficulty level II correct answers	PD_4/PD_1	-0.41	0.009
Difficulty level III correct answers	PD_6/PD_1	-0.38	0.017
Total score correct answers	PD_7/PD_1	-0.45	0.004

level III' segment, i.e. PD_6/PD_1 ($p < 0.02$), as well as for total MOT performance and normalised average pupil diameter after all MOT segments, i.e. PD_7/PD_1 ($p < 0.01$). These results indicate that ATC candidates were experiencing higher stress, i.e. pupil dilation, with a decrease in their performance on the MOT.

A selected experimental paradigm based on RDP task for analysis of saccadic eye movements and MOT task for analysis of multiple-object tracking requires high visual attention and overall oculomotor performance. Such compact 10-minute paradigm has demonstrated dominantly fatigue effects on oculometric SPV feature of ATC candidates, as well as variance in accuracy of their object-tracking visual attention strategies which depends on the tracking task difficulty. These results represent desirable properties of the paradigm for its fine-tuning and further application and validation within ATC selection processes. MOT performance analyses, particularly on the most challenging task segment that required simultaneous tracking of 3 dots out of 10 points, illustrate that the highest performing ATC candidates exhibit considerably superior object-tracking accuracy than the lowest-performing candidates, since the accuracies range from 20% to 80%. Additionally, the results obtained on correlations of normalised pupil diameter feature PD_7/PD_1 with total MOT performance (Table 1) suggest that ATC candidates experienced higher stress when making more mistakes during MOT, likely due to pressure of competing for an attractive job. Altogether, based on the obtained results, the following three features are applicable in future ATC selection processes: average SPV reduction from the first to the last segment (SPV_{rd}), MOT performance during difficulty level III (MOT_{III}), and normalised pupil diameter PD_7/PD_1. The first two features are also mutually uncorrelated in our sample of ATC candidates, suggesting that each of these features should represent relatively independent source of information regarding candidates in future ATC selection processes.

V. COGNITIVE LOAD ASSESSMENT BASED ON SPEECH FEATURES

Speech production is a process which depends on a complex interaction between central and peripheral nervous system innervating around 90-120 muscles. Brain regions vital for speech productions are related to: Broca's area; Premotor cortex; Wernicke's area; Angular gyrus, etc. [75]. Broca's area, by the premotor cortex, sends nerve impulses to all muscles involved in speech production. The most prominent among them are the vocal cords that will either contract or relax. Simultaneously, the respiratory muscles will compress the lungs, and a current of air will flow over the vocal cords. Following the vocal cords, air flows through the rest of the vocal tract, ending with the lips. The vocal tract can be represented by a system with a specific frequency response that defines the characteristics of the produced sound based on its shape, length, thermal capacity, etc. Specific sounds are a result of various small perturbations in the air fluctuation that produces a unique frequency response for the vocal cord excitation signal coming from the lower portions of the speech production system. These types of perturbations are affected by states like stress, fatigue, cognitive load or physical exertion and are, therefore, highly relevant for our research related to the ATC selection process.

Cognitive load is defined as "the load imposed on one's cognitive system when performing a particular task" [76]. It means that cognitive overload can be considered as a type of stress, according to the definition of stress: "psycho-physiological state characterised by subjective strain, dysfunctional physiological activity and deterioration of performance" [20]. Cognitive overload can be caused by different factors, such as variations of alertness, mental fatigue, mental effort, complexity of the task, attentional variations and drowsiness. In the context of ATC selection process cognitive load and its effects on speech features have already been researched [15]–[18], [19]. In our cognitive load experiment, we apply different versions of Stroop test, which have also been used in the context of selection and training of ATCs, e.g. in the analysis of ATC trainees' cognitive control strategies [77]. ATC trainees exhibited higher Stroop test performance, i.e. reduced cognitive interference, in comparison to the control group.

Most of the research into the effects of cognitive load on speech are related to classification of speaker cognitive load from spoken utterances or identification of speech features that show meaningful variation under varying cognitive load conditions. Some of the earliest papers in the field, like [78] and [79], discuss these effects under experimental conditions, like arithmetic and visual tracking tasks. Cognitive load classification using speech/voice data has emerged lately alongside the more widespread use of machine learning methods in the late 2000s. Most of the papers on the topic follow the same procedure, they present novel handmade speech features as extensions of basic speech analysis feature sets, use them for classification of cognitive load in a particular dataset, and finally present the improvements in classifier scores when using the presented features. Some of the first papers that used such a framework [80], [81] present Gaussian Mixture Modelling of cognitive load of reading comprehension tasks and Stroop task variations using basic prosodic and cepstral features. These results were later extended by additional speech features, namely: spectral centroid frequency and amplitude [82], cepstral peak prominence and harmonic-to-noise ratio [83] and vowel formant trajectory-based features [84]. More lately, a comparative analysis of various machine learning methods and speech feature sets for classification of cognitive load induced by different Stroop test variations was conducted [85]. Based on the presented literature, our research on speech physiological reactions during ATC selection process is focused on the impact of cognitive-load-induced sympathetic activity on verbal responses to Stroop tasks, from which a comprehensive set of speech features is computed.

A. STIMULI PARADIGMS AND TASKS

The experimental paradigm applied in the ATC selection process was based on three variations of the Stroop test: Batch Stroop tasks, Time-constrained Stroop tasks and Divided attention Stroop tasks. In all three variants, the Stroop tasks used six colours (black, blue, green, yellow, red and orange) and the experimental stimuli were delivered on a computer screen. All the Stroop test variations were based on verbal task answers which enabled us to analyse the cognitive load and its effects on the participants' speech. The experimental paradigm was 7 minutes long and consisted of 5 main parts:

- **Baseline reading task.** This 1-minute block was used to produce sufficient speech/voice data for any required data normalisations or post-paradigm analyses (see the last block of the paradigm). The participants had to read a generic text.
- **Batch Stroop tasks.** The batch variant of the Stroop test lasted for 1 minute and consisted of 6 batches of 10 tasks (each task remained on the computer screen for 10 seconds). In three of the batches, the colour words and the font colours in which they were written were congruent, while in the other three they were incongruent.
- **Time-constrained Stroop tasks.** This block lasted for 1.5 minutes. In this variant of the Stroop test, 90 consecutive tasks were exchanged on the screen in 0.75-1 second intervals, while the congruent-incongruent conditions were randomised.
- **Divided attention Stroop tasks.** This variant of Stroop test lasted for 2.5 minutes and consisted of 6 blocks of 20 tasks. The task schedule was similar to the time-constrained Stroop tasks, but with consecutive tasks exchanging on the screen in 1.25-second intervals. This Stroop test was accompanied by the auditory stimuli counting task. Two types of auditory stimuli (high and low tone) were administered through the headphones, and the participants' task was to count the high tones. After each block of 20 Stroop tasks, the participants were required to say the number of high tones they counted in the current block.
- **Reading task at the end.** At the end of the paradigm, the participants had to read the same text from the beginning of the paradigm. This allowed for the analysis of the effects of the experimental paradigm on the participants' speech features.

On the basis of these tasks, we can estimate performance-based measures of perceptual and processing speed, cognitive inhibition, and divided and selective attention (multi-tasking capabilities), which are generally relevant for the ATC profession [4]. All the Stroop test variations were based on verbal task answers which additionally enables us to perform speech-based analyses of individual mental states like cognitive load/overload. The Stroop task was used in such a context in a number of previous papers, both in ATC samples and military applications [16], [17], [19], [83], [84], [86], [87].

The task answers were recorded using the built-in headset microphone (Sennheiser PC 360 G4ME). Visualisation of the whole speech-based Stroop paradigm can be seen in Fig. 7.

B. MEASUREMENTS AND RESULTS

After the design of an appropriate stimulation paradigm (Fig. 7) and ATC candidate data collection, speech-based analysis includes the following steps: segmentation of answer utterances; extraction of Low-Level-Descriptive signals (LLDs) from the segmented utterances; computation of statistical features over the extracted LLDs of segmented utterances, e.g. 10th and 90th centiles, arithmetic mean, median, standard deviation, minimum and maximum etc.; labelling the computed statistical features; performing the appropriate statistical analysis and inference.

LLDs are continuous signals computed/estimated by using signal processing techniques (spectral and cepstral analysis, autocorrelation analysis, etc.) that aim to describe vocal source and vocal tract behaviour for specific recorded speech utterances. For example, F0 is an estimate of the base harmonic that the vibrating vocal cords are producing during vocalisation; the RMS is a signal-processing-based estimate of the energy of the sound pressure produced at the lips during vocalisation and recorded by the microphone. LLDs like the formant frequencies (F1-F4) or Mel Frequency Cepstral Coefficients (MFCCs) aim to describe the spectral and cepstral behaviour of the recorded utterances. All LLD extraction was performed by using the openSMILE feature extraction tool [88] which is a standard research and industry tool. The system was used in a configuration with the Extended Geneva Minimalistic Acoustic Parameter Set for Voice Research and Affective Computing (eGeMAPS) [89] which is also a high-performing standard set of LLDs, widely used in affective research of speech.

To identify the features that significantly discriminate between the two Stroop conditions (congruent vs. incongruent), we have used linear mixed effects models (LMEM), which were used in our previous research on speech features of acoustic startle responses [32]. The LMEMs enabled us to model individual feature variations over the two Stroop conditions, taking into account the three Stroop task types, the sex of the participants and inter-subject variability. The fitted LMEM models were then tested using the F test for all features that exhibit statistically significant differences ($p<0.01$) between the two Stroop conditions.

A total of 26 features exhibited statistically significant differences between the two Stroop conditions. Among the identified features, three groups are most prominent:

- features related to voice intensity, e.g. statistic functionals of voice loudness (perceived RMS voice energy) like loudness mean and median, 20th and 80th percentiles and several others
- features related to the spectral characteristic of speech, e.g. F0, formant amplitudes (prominent peaks in the spectral characteristic) in relation to F0, spectral flux, MFCCs, jitter, etc.

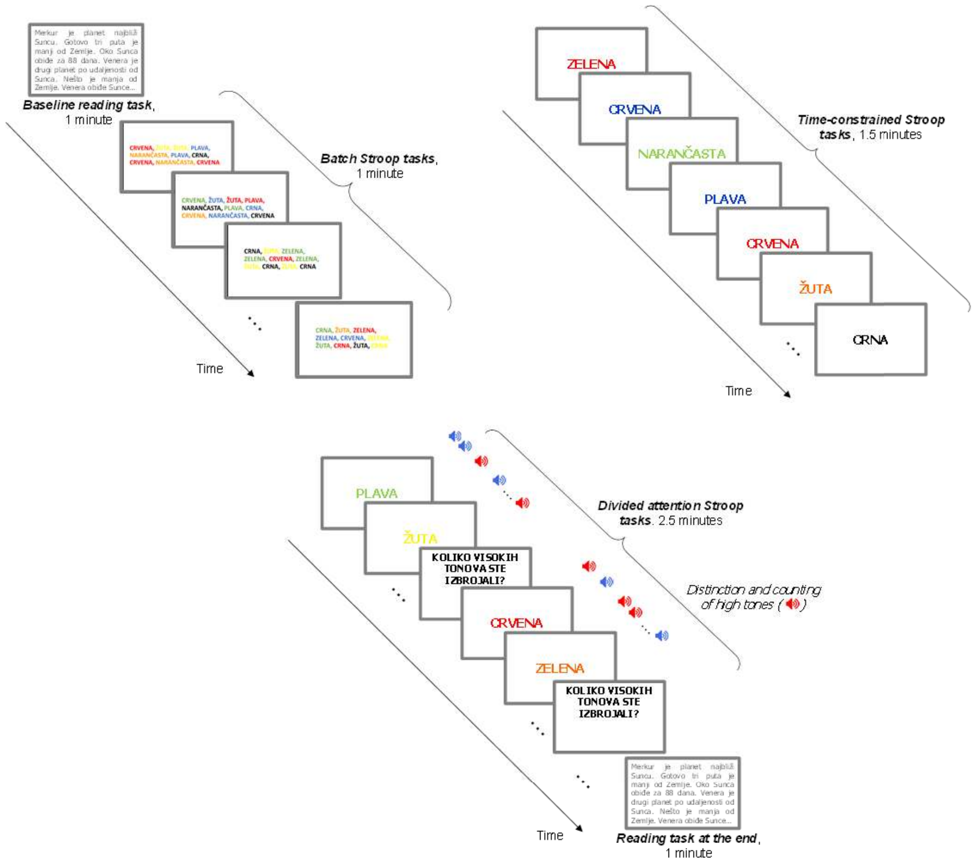

FIGURE 7. **Visualisation of the ATC cognitive load Stroop-based paradigm with all its main parts. The colour words and instructions were written in Croatian.**

- prosodic features related to speech rate/speed, e.g. number of voiced segments per second and mean voiced/unvoiced segment lengths in seconds.

A significant proportion of the identified features is already known to be related to sympathetic activity [16], [17], [19], [86], [87], [83], [84]. For example, F0 and loudness are expected to rise in the cases of heightened sympathetic activity [31], with changes in the whole spectral composition ''shifting'' to higher values.

To illustrate the inter-candidate variance in the speech-based measure of induced cognitive load, we have chosen loudness_mean, due to its ease of interpretability and the consistent presence of various statistical descriptives of the loudness LLD in the final set of 26 features. The induced cognitive load for each candidate was represented by the loudness_mean difference (LMD) feature: calculated as the difference between the average value of loudness_mean across the incongruent Stroop condition (higher cognitive load) and the average value of loudness_mean across the congruent Stroop condition (lower cognitive load). The corresponding distribution (histogram) is shown in Fig. 8.

The psychometric Stroop test performance-based features extracted from the Time-constrained Stroop tasks were: the average reaction time for the congruent condition (directly

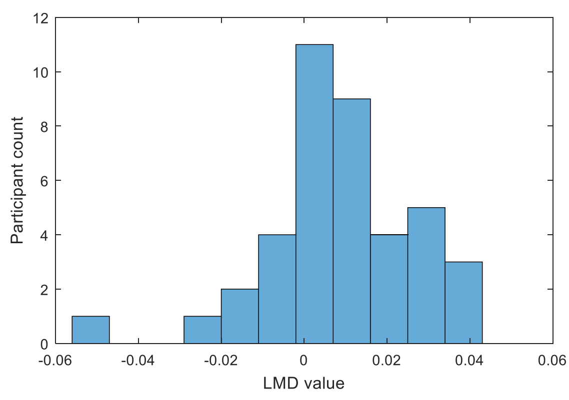

FIGURE 8. **Distribution of the LMD measure of sympathetic arousal induced by cognitive load.**

measures perceptive speed); the average reaction time for the incongruent Stroop tasks (measures both perceptive speed and cognitive inhibition); the difference in average reaction times between the incongruent and congruent Stroop tasks (measures cognitive inhibition); and relative error rate feature ST_{tc}: ratio of incorrect responses vs. total number of individual Stroop tasks. The psychometric Stroop test performance-based feature ST_{da} extracted from the Divided

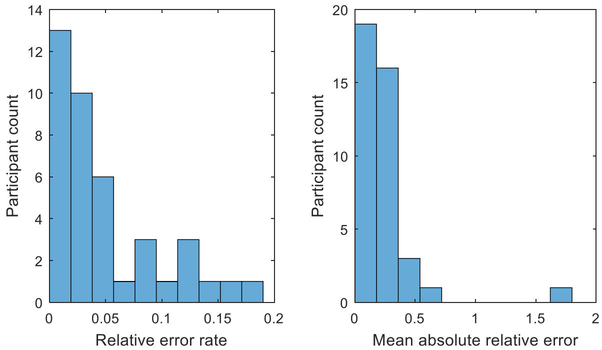

FIGURE 9. Distributions of task performance metrics for the Time-constrained Stroop tasks (ST_{tc}, left) and the high tone counting during Divided attention Stroop tasks (ST_{da}, right).

attention Stroop tasks was a mean absolute relative error of the candidates' responses in the high tone counting task. Illustration of task evaluation metrics for both the Divided attention Stroop tasks and Time-constrained Stroop tasks can be seen in Fig. 9 which shows that the feature distributions are particularly suited for identification of low-performing candidates, due to the pronounced positive skewness in both distributions.

From the set of speech features which were extracted, ranking of the candidates can be even intuitively interpreted, e.g., for a normally distributed sample of response times to Stroop tasks, participants with lower response times can be regarded as superior to those with higher response times. In the context of speech features responses to cognitive overload as a type of stress [16], [19], [86], [87], participants with lower LMD are exhibiting either lower sympathetic responses, or higher parasympathetic responses, or better balance of autonomic nervous system. Such individuals would be better suited for inherently highly stressful jobs, like pilots, ATCs etc.

Finally, we have to stress that speech data are more readily available than peripheral physiology since many practical applications in aviation already use speech communication systems. Therefore, we do believe that speech research in the context of selection process may have significant potential.

VI. DISCUSSION

High performance of ATC candidates requires extraordinary cognitive and emotional human potential, which might be assessed by proposed clusters of different multimodal features presented in this article. In this paper, we selected a set of 3 × 3 multimodal physiological, oculometric and speech features as tools and means for ATC candidates' performance evaluation and selection. Radar chart in Fig. 10 illustrates candidates' variability within multimodal and multidimensional performance space, which is spanned by 9 selected axes for evaluation of ATC candidates, related to stress resilience, visual attention and fatigue, as well as cognitive load. 3 × 3 multimodal radar chart includes: 3 features, RSA, SR and CA, for stress resilience assessment based on physiological measurements; 3 features, SPV_{rd}, PD_7/PD_1, and MOT_{III} for visual attention and fatigue assessment based on oculometric measurements; 3 features, LMD, ST_{tc}, and ST_{da}, for cognitive load assessment based on speech measurements. Thick line connects medians on each axis, shaded area represents interquartile range of each feature, while the endpoints of each axis correspond to the minimum and maximum value of the respective feature in our sample of ATC candidates. Outer values of each axis on the radar chart are associated with better ATC performance. Proposed 3 × 3 multimodal metrics integrated with FEAST/DART & AT-SAT can be used as enhancement of standardised ATC selection process. Such integration can minimise the attrition/drop-out rate of selected ATC candidates in future selection processes. Specifically, better performing ATC candidates should have higher RSA and CA features, lower SR, SPV_{rd}, PD_7/PD_1 and LMD features, higher MOT_{III} scoring and lower relative errors ST_{tc} and ST_{da} than the rest of the group. This was the first time, to the best of our knowledge, that such integrated multimodal psychophysiological assessment has been applied in the ATC selection process. Applied stimulation paradigms are low-cost and relatively short, less than 45 minutes. Proposed multidimensional ATC performance space emphasises the importance of combining different multimodal features in enhancing performance predictive power, i.e. any single feature in assessment of complex human behavioural performance is a relatively weak predictor, what underscores the need to carefully combine multimodal features in the ATC selection process. Therefore, the need to combine different multimodal features is logical step forward in enhancement of ATC selection protocols, in order to minimise relatively high attrition rates during the training period before full ATC qualification, which can reach and exceed 30% in both civilian and military settings [3], [90], [91]. Proposed multimodal and multidimensional ATC performance space and its extensions with the most pertinent bio-neuro-psycho-social features should be an important step forward to comprehensive multidisciplinary selection process, which has to take into account discriminative power of each feature and its cost. Prospective research based on machine learning and data-driven integration of various features, which has potential to discriminate high- vs. low-performing ATC candidates, deserves additional efforts. Features selection and classification based on machine learning, as opposed to statistical methods, that are mostly used in related work, would explore more complex interactions between various features in a highly non-linear manner associated with the outcome performance. Additionally, how we label data needs particular attention, like fuzzy logic labelling along hierarchical data chain, including: stress resilience (''high'', ''medium'', ''low''), visual attention (''high'', ''medium'', ''low'') and cognitive capacities (''high'', ''medium'', ''low''), as well as overall on-the-job performance (''high'', ''medium'', ''low''). This labelling process requires longitudinal prospective research years after the initial measurement but could provide highly valuable insight into the predictive power of the proposed approach. Proposed physiological, oculometric and speech features

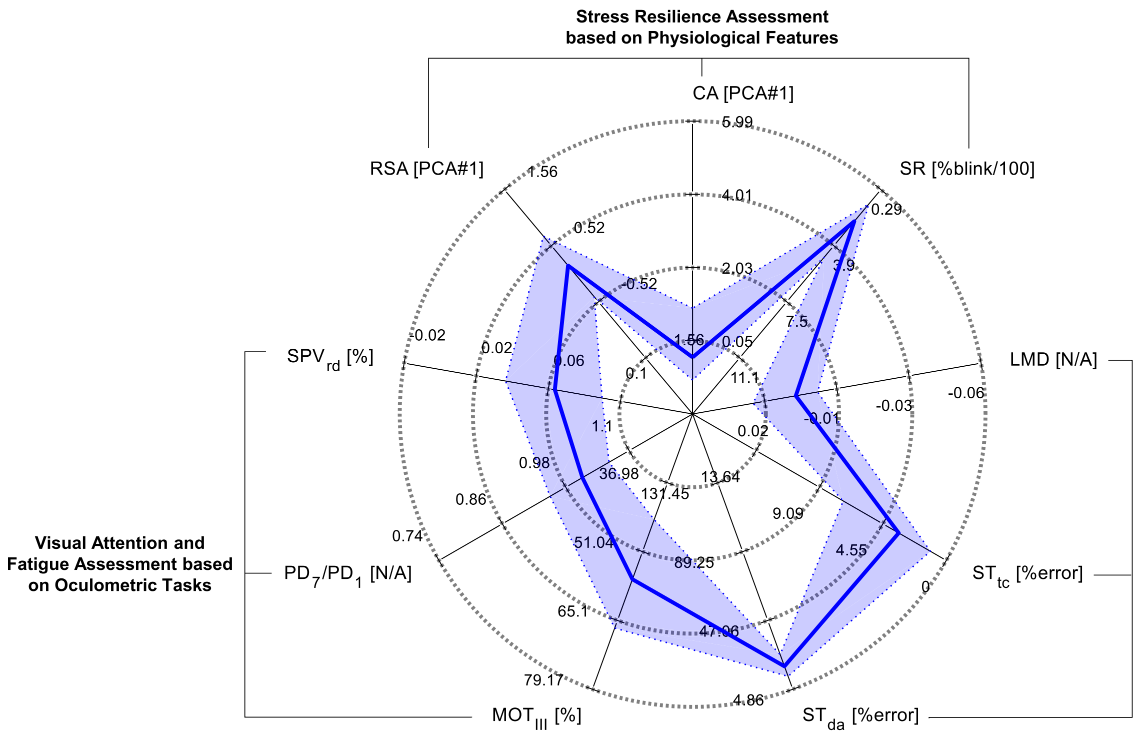

FIGURE 10. 3 × 3 multimodal performance assessment radar chart.

could be combined with different psychological questionnaires, like The Connor-Davidson Resilience scale (CD-RISC) [92], State-Trait Anxiety Inventory (STAI) [93], Beck Depression Inventory (BDI) [94] etc., as well as domain expert assessments, to enhance the quality of ATC selection process.

Finally, the presented concept of technologically assisted ATC selection process using multimodal features during different experimental conditions might also strengthen the quality of future selection process of other stressful occupations like first responders, civilian and military pilots, astronauts and military personnel.

ACKNOWLEDGMENT

The authors particularly thank the director of the Croatian Air Traffic Control – Croatia Control, Mr. V. Bagarić, for project leadership, as well as his associates for their assistance in experimental measurements. The authors also thank anonymous reviewers for their valuable comments.

REFERENCES

[1] Eurocontrol. *Frequently Asked Questions: Is it Really Stressful, Being an Air Traffic Controller?* Accessed: Nov. 26, 2019. [Online]. Available: https://atco.eurocontrol.int/

[2] G. Costa, "Occupational stress and stress prevention in air traffic control," Int. Labour Org., Geneva, Switzerland, Work. Paper CONDFT/WP.6/1995. Accessed: Oct. 18, 2019. [Online]. Available: https://www.ilo.org/wcmsp5/groups/public/—ed_protect/—protrav/—safework/documents/publication/wcms_250120.pdf

[3] N. De Koninck, "Reaching out to the next generation: Student air traffic controller selection & Training process," World ATM Congr., Madrid, Spain, Mar. 2017. [Online]. Available: https://docplayer.net/55311464-Reaching-out-to-the-next-generation.html

[4] A. Nunes and A. F. Kramer, "Experience-based mitigation of age-related performance declines: Evidence from air traffic control," *J. Exp. Psychol. Appl.*, vol. 15, no. 1, pp. 12–24, 2009.

[5] A. L. Scarborough and M. K. Bleckley, "Air traffic scenarios test: Will modifying the instructions change performance?" Federal Aviation, Washington, DC, USA, Tech. Rep. 20591.DOT/FAA/AM-15/7, Apr. 2015. [Online]. Available: https://libraryonline.erau.edu/online-full-text/faa-aviation-medicine-reports/AM15-07.pdf

[6] P. Aricò, G. Borghini, G. Di Flumeri, S. Bonelli, A. Golfetti, I. Graziani, S. Pozzi, J.-P. Imbert, G. Granger, R. Benhacene, D. Schaefer, and F. Babiloni, "Human factors and neurophysiological metrics in air traffic control: A critical review," *IEEE Rev. Biomed. Eng.*, vol. 10, pp. 250–263, Apr. 2017.

[7] E. H. Magnusdottir, K. R. Johannsdottir, C. Bean, B. Olafsson, and J. Gudnason, "Cognitive workload classification using cardiovascular measures and dynamic features," in *Proc. 8th IEEE Int. Conf. Cognit. Infocommun. (CogInfoCom)*, Sep. 2017, pp. 351–356.

[8] F. G. Paas and J. J. G. van Merriënboer, "Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach," *J. Educ. Psychol.*, vol. 86, no. 1, pp. 122–133, 1994.

[9] P. W. Van Gerven, F. Paas, J. J. van Merriënboer, and H. G. Schmidt, "Memory load and the cognitive pupillary response in aging," *Psychophysiology*, vol. 41, no. 2, pp. 167–174, Mar. 2004.

[10] E. Abdulin and O. Komogortsev, "User eye fatigue detection via eye movement behaviour," in *Proc. 33rd Annu. ACM Conf. Extended Abstracts Hum. Factors Comput. Syst.*, Apr. 2015, pp. 1265–1270.

[11] H. S. Meyerhoff, F. Papenmeier, and M. Huff, "Studying visual attention using the multiple object tracking paradigm: A tutorial review," *Atten. Percept. Psychophys.*, vol. 79, no. 5, pp. 1255–1274, Jul. 2017.

[12] 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. Int.*, vol. 30, no. 3, pp. 220–236, Jan. 2014.

[13] H. P. Greeley, E. Friets, J. P. Wilson, S. Raghavan, J. Picone, and J. Berg, "Detecting fatigue from voice using speech recognition," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol.*, Aug. 2006, pp. 567–571.

[14] H. P. Greeley, J. Berg, E. Friets, J. Wilson, G. Greenough, J. Picone, J. Whitmore, and T. Nesthus, "Fatigue estimation using voice analysis," *Behav. Res. Methods*, vol. 39, pp. 610–619, Aug. 2007.

[15] M. Hagmüller, G. Kubin, and E. Rank, *Evaluation of the Human Voice for Indications of Workload Induced Stress in the Aviation Environment*, document EEC Note 18/06, INO-1 AC STSS Project, Eurocontrol, Dec. 2006.
[16] K. H. Huttunen, H. I. Keränen, R. J. Pääkkönen, R. P. Eskelinen-Rönkä, and T. K. Leino, "Effect of cognitive load on articulation rate and formant frequencies during simulator flights," *J. Acoust. Soc. Amer.*, vol. 129, no. 3, pp. 1580–1593, Mar. 2011.
[17] H. I. Keränen, R. J. Pääkkönen, R. P. Eskelinen-Rönkä, and T. K. Leino, "Effect of cognitive load on speech prosody in aviation: Evidence from military simulator flights," *Appl. Ergon.*, vol. 42, no. 2, pp. 348–357, Jan. 2011.
[18] J. Luig and A. Sontacchi, "Workload monitoring through speech analysis: Towards a system for air traffic control," in *Proc. 27th Congr. Int. Council Aeronaut. Sci.*, Sep. 2010, pp. 1–10, Paper ICAS 2010-10.7.4. [Online]. Available: http://www.icas.org/ICAS_ARCHIVE/ICAS2010/ABSTRACTS/669.HTM
[19] E. H. Magnusdottir, M. Borsky, M. Meier, K. Johannsdottir, and J. Gudnason, "Monitoring cognitive workload using vocal tract and voice source features," *Periodica Polytechnica Elect. Eng. Comput. Sci.*, vol. 61, no. 4, pp. 297–304, May 2017.
[20] I. R. Murray, C. Baber, and A. South, "Towards a definition and working model of stress and its effects on speech," *Speech Commun.*, vol. 20, nos. 1–2, pp. 3–12, Nov. 1996.
[21] J. Whitmore and S. Fisher, "Speech during sustained operations," *Speech Commun.*, vol. 20, pp. 55–70, Nov. 1996.
[22] D. E. Glenn, D. T. Acheson, M. A. Geyer, C. M. Nievergelt, D. G. Baker, V. B. Risbrough, and MRS Team, "High and low threshold for startle reactivity associated with PTSD symptoms but not PTSD risk: Evidence from a prospective study of active duty marines," *Depression Anxiety*, vol. 33, no. 3, pp. 192–202, Feb. 2016.
[23] F. R. Walker, K. Pfingst, L. Carnevali, A. Sgoifo, and E. Nalivaiko, "In the search for integrative biomarker of resilience to psychological stress," *Neurosci. Biobehav. Rev.*, vol. 74, pp. 310–320, Mar. 2017.
[24] J. R. Stroop, "Studies of interference in serial verbal reactions," *J. Exp. Psychol.*, vol. 18, no. 6, pp. 643–662, Dec. 1935.
[25] C. M. Armstrong, G. M. Reger, J. Edwards, A. A. Rizzo, C. G. Courtney, and T. D. Parsons, "Validity of the virtual reality Stroop task (VRST) in active duty military," *J. Clin. Exp. Neuropsychol.*, vol. 35, no. 2, pp. 113–123, 2013.
[26] J. M. G. Williams, A. Mathews, and C. MacLeod, "The emotional Stroop task and psychopathology," *Psychol. Bull.*, vol. 120, no. 1, pp. 3–24, Jul. 1996.
[27] K. Ćosić, S. Popović, D. Kukolja, M. Horvat, and B. Dropuljić, "Physiology-driven adaptive virtual reality stimulation for prevention and treatment of stress related disorders," *Cyberpsychol., Behav., Social Netw.*, vol. 13, no. 1, pp. 73–78, Feb. 2010.
[28] K. Ćosić, S. Popović, M. Horvat, D. Kukolja, B. Dropuljić, I. Kostović, M. Judaš, M. Radoš, M. Radoš, L. Vasung, B. B. Spajić, S. Doričić, and D. Mesić, "Virtual reality adaptive stimulation in stress resistance training," in *Proc. Mental Health Well-Being Across Mil. Spectr.* Neuilly-sur-Seine Cedex, France: NATO Research and Technology Organisation, Apr. 2011, pp. 4-1–4-18, Paper RTO-MP-HFM-205. [Online]. Available: https://pdfs.semanticscholar.org/942d/22e84c84aeee688f500647f525ec2368749c.pdf
[29] K. Ćosić, S. Popović, M. Horvat, B. Dropuljić, B. Kovač, and M. Jakovljević, "Computer-aided psychotherapy based on multimodal elicitation, estimation and regulation of emotion," *Psychiatria Danubina*, vol. 25, no. 3, pp. 340–346, 2013.
[30] K. Cosić, S. Popović, M. Horvat, D. Kukolja, B. Dropuljić, B. Kovać, and I. Fabek, "Multimodal paradigm for mental readiness training and PTSD prevention," in *New Tools to Enhance Posttraumatic Stress Disorder Diagnosis Treatment: Invisible Wounds War*, B. K. Wiederhold, Ed. Amsterdam, The Netherlands: IOS Press, 2013, pp. 102–116.
[31] K. Ćosić, S. Popović, D. Kukolja, B. Dropuljić, D. Ivanec, and M. Tonković, "Multimodal analysis of startle type responses," *Comput. Methods Programs Biomed.*, vol. 129, pp. 186–202, Jun. 2016.
[32] B. Dropuljić, I. Mijić, D. Petrinović, T. Jovanovic, and K. Ćosić, "Vocal analysis of acoustic startle responses," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 2, pp. 318–329, Nov. 2017.
[33] D. Kukolja, S. Popović, M. Horvat, B. Kovač, and K. Ćosić, "Comparative analysis of emotion estimation methods based on physiological measurements for real-time applications," *Int. J. Hum.-Comput. Stud.*, vol. 72, nos. 10–11, pp. 717–727, Oct. 2014.
[34] S. Popović, M. Horvat, D. Kukolja, B. Dropuljić, and K. Ćosić, "Stress inoculation training supported by physiology-driven adaptive virtual reality stimulation," in *Proc. Annu. Rev. Cybertherapy Telemed.*, 2009, pp. 50–54.
[35] U. Ahlstrom and F. J. Friedman-Berg, "Using eye movement activity as a correlate of cognitive workload," *Int. J. Ind. Ergonom.*, vol. 36, no. 7, pp. 623–636, Jul. 2006.
[36] J. B. Brookings, G. F. Wilson, and C. R. Swain, "Psychophysiological responses to changes in workload during simulated air traffic control," *Biol. Psychol.*, vol. 42, no. 3, pp. 361–377, Feb. 1996.
[37] D. C. Evans and M. Fendley, "A multi-measure approach for connecting cognitive workload and automation," *Int. J. Hum.-Comput. Stud.*, vol. 97, pp. 182–189, Jan. 2017.
[38] M. Marchitto, S. Benedetto, T. Baccino, and J. J. Cañas, "Air traffic control: Ocular metrics reflect cognitive complexity," *Int. J. Ind. Ergonom.*, vol. 54, pp. 120–130, Jul. 2016.
[39] I. Mijić, M. Šarlija, and D. Petrinovìc, "MMOD-COG: A database for multimodal cognitive load classification," in *Proc. 11th Int. Symp. Image Signal Process. Anal.*, Sep. 2019, pp. 15–20.
[40] S. Rodríguez, L. Sánchez, P. López, and J. J. Cañas, "Pupillometry to assess air traffic controller workload through the mental workload model," in *Proc. 5th Int. Conf. Appl. Theory Autom. Command Control Syst.*, Sep. 2015, pp. 95–104.
[41] M. Vukovic, V. Sethu, J. Parker, L. Cavedon, M. Lech, and J. Thangarajah, "Estimating cognitive load from speech gathered in a complex real-life training exercise," *Int. J. Hum.-Comput. Stud.*, vol. 124, pp. 116–133, Apr. 2019.
[42] G. E. Strangman, V. Ivkovic, and Q. Zhang, "Wearable brain imaging with multimodal physiological monitoring," *J. Appl. Physiol.*, vol. 124, no. 3, pp. 564–572, Jul. 2017.
[43] K. Ćosić, S. Popović, B. Kovać, D. Kukolja, D. Ivanec, and T. Jovanovic, "System for evaluation of cognitive performance under the emotional stressors," in *Proc. 6th Int. Conf. Adv. Cogn. Technol. Appl. (COGNITIVE)*, H. Lounis and D. Josyula, Eds. Wilmington, DE, USA: IARIA, 2014, pp. 239–245.
[44] K. Ćosić, M. Šarlija, V. Ivkovic, Q. Zhang, G. Strangman, and S. Popović, "Stress resilience assessment based on physiological features in selection of air traffic controllers," *IEEE Access*, vol. 7, pp. 41989–42005, 2019.
[45] D. Fletcher and M. Sarkar, "Psychological resilience: A review and critique of definitions, concepts, and theory," *Eur. Psychol.*, vol. 18, no. 1, pp. 12–23, Apr. 2013.
[46] S. J. Russo, J. W. Murrough, M.-H. Han, D. S. Charney, and E. J. Nestler, "Neurobiology of resilience," *Nature Neurosci.*, vol. 15, no. 11, pp. 1475–1484, Nov. 2012.
[47] M. Greener, "For everything there is a season: Biological rhythms in psychiatry," *Prog. Neurol. Psychiatry*, vol. 21, no. 4, pp. 30–32, Nov. 2017.
[48] R.-P. Juster, B. S. McEwen, and S. J. Lupien, "Allostatic load biomarkers of chronic stress and impact on health and cognition," *Neurosci. Biobehav. Rev.*, vol. 35, no. 1, pp. 2–16, Sep. 2010.
[49] M. K. Taylor, K. P. Sausen, E. G. Potterat, L. R. Mujica-Parodi, J. P. Reis, A. E. Markham, G. A. Padilla, and D. L. Taylor, "Stressful military training: Endocrine reactivity, performance, and psychological impact," *Aviation, Space, Environ. Med.*, vol. 78, no. 12, pp. 1143–1149, Dec. 2007.
[50] P. Grossman and E. W. Taylor, "Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions," *Biol. Psychol.*, vol. 74, no. 2, pp. 263–285, Feb. 2007.
[51] D. E. Bradford, J. T. Kaye, and J. J. Curtin, "Not just noise: Individual differences in general startle reactivity predict startle response to uncertain and certain threat," *Psychophysiology*, vol. 51, no. 5, pp. 407–411, Feb. 2014.
[52] G. G. L. Souza, L. N. Magalhaes, T. A. R. Da Cruz, A. C. F. Mendonça-De-Souza, A. F. A. Duarte, N. L. Fischer, W. F. Souza, E. D. S. F. Coutinho, J. Vila, S. Gleiser, and I. Figueira, "Resting vagal control and resilience as predictors of cardiovascular allostasis in peacekeepers," *Stress*, vol. 16, no. 4, pp. 377–383, Feb. 2013.
[53] M. E. Goldberg and M. F. Walker, "The control of gaze," in *Principles of Neural Science*, 5th ed. New York, NY, USA: McGraw-Hill, 2013, pp. 894–916.
[54] J. F. Hejtmancik, P. Cabrera, Y. Chen, O. M'Hamdi, and J. M. Nickerson, "Vision," in *Conn's Translational Neuroscience*. Amsterdam, The Netherlands: Elsevier, 2017, pp. 399–438.
[55] L. L. Di Stasi, A. Catena, J. J. Cañas, S. L. Macknik, and S. Martinez-Conde, "Saccadic velocity as an arousal index in naturalistic tasks," *Neurosci. Biobehav. Rev.*, vol. 37, no. 5, pp. 968–975, Jun. 2013.

[56] L. L. Di Stasi, M. Marchitto, A. Antolí, and J. J. Cañas, "Saccadic peak velocity as an alternative index of operator attention: A short review," *Eur. Rev. Appl. Psychol.*, vol. 63, no. 6, pp. 335–343, Nov. 2013.

[57] M. Causse, F. Lancelot, J. Maillant, J. Behrend, M. Cousy, and N. Schneider, "Encoding decisions and expertise in the operator's eyes: Using eye-tracking as input for system adaptation," *Int. J. Hum.-Comput. Stud.*, vol. 125, pp. 55–65, May 2019.

[58] G. Hauland, "Measuring individual and team situation awareness during planning tasks in training of en route air traffic control," *Int. J. Aviation Psychol.*, vol. 18, no. 3, pp. 290–304, Jun. 2008.

[59] K. Moore and L. Gugerty, "Development of a novel measure of situation awareness: The case for eye movement analysis," in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*. Los Angeles, CA, USA: Sage, vol. 54, no. 19, Sep. 2010, pp. 1650–1654.

[60] Y. Wang, W. Cong, B. Dong, F. Wu, and M. Hu, "Statistical analysis of air traffic controllers' eye movements," in *Proc. 11th USA/Eur. ATM R&D Seminar*, 2015, pp. 1–9.

[61] C. Yu, E. M.-Y. Wang, W.-C. Li, and G. Braithwaite, "Pilots' visual scan patterns and situation awareness in flight operations," *Aviation, Space, Environ. Med.*, vol. 85, no. 7, pp. 708–714, Jul. 2014.

[62] M. K. Eckstein, B. Guerra-Carrillo, A. T. M. Singley, and S. A. Bunge, "Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development?" *Develop. Cognit. Neurosci.*, vol. 25, pp. 69–91, Jun. 2017.

[63] H. A. Slagter, K. Georgopoulou, and M. J. Frank, "Spontaneous eye blink rate predicts learning from negative, but not positive, outcomes," *Neuropsychologia*, vol. 71, pp. 126–132, May 2015.

[64] J. E. Thropp, J. F. V. Scallon, and P. Buza, "PERCLOS as an indicator of slow-onset hypoxia in aviation," *Aerosp. Med. Hum. Perform.*, vol. 89, no. 8, pp. 700–707, Aug. 2018.

[65] S. Joshi, Y. Li, R. M. Kalwani, and J. I. Gold, "Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex," *Neuron*, vol. 89, no. 1, pp. 221–234, Jan. 2016.

[66] R. S. Larsen and J. Waters, "Neuromodulatory correlates of pupil dilation," *Frontiers Neural Circuits*, vol. 12, Mar. 2018, Art. no. 21.

[67] O. Borodovitsyna, N. Joshi, and D. Chandler, "Persistent stress-induced neuroplastic changes in the locus coeruleus/norepinephrine system," *Neural Plasticity*, vol. 2018, Jun. 2018, Art. no. 1892570.

[68] E. R. Samuels and E. Szabadi, "Functional neuroanatomy of the noradrenergic locus coeruleus: Its roles in the regulation of arousal and autonomic function part I: Principles of functional organisation," *Current Neuropharmacol.*, vol. 6, no. 3, pp. 235–253, Sep. 2008.

[69] I. Sluganovic, M. Roeschlin, K. B. Rasmussen, and I. Martinovic, "Using reflexive eye movements for fast challenge-response authentication," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Oct. 2016, pp. 1056–1067.

[70] H. M. Fehd and A. E. Seiffert, "Eye movements during multiple object tracking: Where do participants look?" *Cognition*, vol. 108, no. 1, pp. 201–209, Jul. 2008.

[71] A. Tran and J. E. Hoffman, "Visual attention is required for multiple object tracking," *J. Exp. Psychol., Hum. Perception Perform.*, vol. 42, no. 12, pp. 2103–2114, Dec. 2016.

[72] G. J. Zelinsky and M. B. Neider, "An eye movement analysis of multiple object tracking in a realistic environment," *Vis. Cogn.*, vol. 16, no. 5, pp. 553–566, Jun. 2008.

[73] R. Allen, P. McGeorge, D. Pearson, and A. B. Milne, "Attention and expertise in multiple target tracking," *Appl. Cogn. Psychol., Off. J. Soc. Appl. Res. Memory Cognition*, vol. 18, no. 3, pp. 337–347, Mar. 2004.

[74] R. Walker, D. G. Walker, M. Husain, and C. Kennard, "Control of voluntary and reflexive saccades," *Exp. Brain Res.*, vol. 130, no. 4, pp. 540–544, Feb. 2000.

[75] R. M. Lazar and J. P. Mohr, "Revisiting the contributions of Paul Broca to the study of aphasia," *Neuropsychol. Rev.*, vol. 21, no. 3, pp. 236–239, Sep. 2011.

[76] F. Paas, J. E. Tuovinen, H. Tabbers, and P. W. V. Gerven, "Cognitive load measurement as a means to advance cognitive load theory," *Educ. Psychol.*, vol. 38, no. 1, pp. 63–71, Mar. 2003.

[77] S. Arbula, M. Capizzi, N. Lombardo, and A. Vallesi, "How life experience shapes cognitive control strategies: The case of air traffic control training," *PLoS ONE*, vol. 11, no. 6, Jun. 2016, Art. no. e0157731.

[78] M. H. L. Hecker, K. N. Stevens, G. von Bismarck, and C. E. Williams, "Manifestations of task-induced stress in the acoustic speech signal," *J. Acoust. Soc. Amer.*, vol. 44, no. 4, pp. 993–1001, Jul. 1968.

[79] S. E. Lively, D. B. Pisoni, W. Van Summers, and R. H. Bernacki, "Effects of cognitive workload on speech production: Acoustic analyses and perceptual consequences," *J. Acoust. Soc. Amer.*, vol. 93, no. 5, pp. 2962–2973, Jan. 1993.

[80] B. Yin, F. Chen, N. Ruiz, and E. Ambikairajah, "Speech-based cognitive load monitoring system," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2008, pp. 2041–2044.

[81] B. Yin, N. Ruiz, F. Chen, and M. A. Khawaja, "Automatic cognitive load detection from speech features," in *Proc. 19th Australasian Conf. Comput.-Hum. Interact., Entertaining User Inter.*, Nov. 2007, pp. 249–255.

[82] P. N. Le, E. Ambikairajah, J. Epps, V. Sethu, and E. H. Choi, "Investigation of spectral centroid features for cognitive load classification," *Speech Commun.*, vol. 53, no. 4, pp. 540–551, Apr. 2011.

[83] T. F. Yap, J. Epps, E. Ambikairajah, and E. H. Choi, "Voice source features for cognitive load classification," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 5700–5703.

[84] T. F. Yap, J. Epps, E. Ambikairajah, and E. H. Choi, "Formant frequencies under cognitive load: Effects and classification," *EURASIP J. Adv. Signal Process.*, vol. 2011, no. 1, Dec. 2011, Art. no. 219253.

[85] M. V. Segbroeck, R. Travadi, C. Vaz, J. Kim, M. P. Black, A. Potamianos, and S. S. Narayanan, "Classification of cognitive load from speech using an i-vector framework," in *Proc. 15th Annu. Conf. Int. Speech Commun. Assoc.*, Sep. 2014, pp. 751–755.

[86] M. K. MacPherson, D. Abur, and C. E. Stepp, "Acoustic measures of voice and physiologic measures of autonomic arousal during speech as a function of cognitive load," *J. Voice*, vol. 31, no. 4, pp. 504–512, Jul. 2017.

[87] K. R. Scherer, D. Grandjean, T. Johnstone, G. Klasmeyer, and T. Bänziger, "Acoustic correlates of task load and stress," in *Proc. 7th Int. Conf. Spoken Lang. Process.*, Sep. 2002, pp. 2017–2020.

[88] F. Eyben, F. Weninger, F. Gross, and B. Schuller, "Recent developments in opensmile, the munich open-source multimedia feature extractor," in *Proc. 21st ACM Int. Conf. Multimedia*, Oct. 2013, pp. 835–838.

[89] F. Eyben, K. R. Scherer, B. W. Schuller, J. Sundberg, E. André, C. Busso, L. Y. Devillers, J. Epps, P. Laukka, S. S. Narayanan, K. P. Truong, "The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing," *IEEE Trans. Affect. Comput.*, vol. 7, no. 2, pp. 190–202, Apr./Jun. 2015.

[90] D. Broach, C. L. Byrne, C. A. Manning, L. Pierce, D. McCauley, and M. K. Bleckley, "The validity of the air traffic selection and training (AT-SAT) test battery in operational use," Federal Aviation Admin., Civil Aerosp. Med. Inst., Oklahoma City, OK, USA, Tech. Rep. DOT/FAA/AM-13/3, Mar. 2013.

[91] T. R. Carretta and R. E. King, "Improved military air traffic controller selection methods as measured by subsequent training performance," *Aviation Space, Environ. Med.*, vol. 79, no. 1, pp. 36–43, Jan. 2008.

[92] K. M. Connor and J. R. Davidson, "Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC)," *Depression Anxiety*, vol. 18, no. 2, pp. 76–82, Sep. 2003.

[93] T. M. Marteau and H. Bekker, "The development of a six-item short-form of the state scale of the Spielberger state—Trait anxiety inventory (STAI)," *Brit. J. Clin. Psychol.*, vol. 31, no. 3, pp. 301–306, Sep. 1992.

[94] A. T. Beck, C. Ward, M. Mendelson, J. Mock, and J. Erbaugh, "Beck depression inventory (BDI)," *Arch. Gen. Psychiatry*, vol. 4, no. 6, pp. 561–571, 1961.

KRE IMIR OSI received the Dipl.Ing., M.S., and Ph.D. degrees in electrical engineering from the Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, in 1973, 1978, and 1984, respectively. He is currently a Professor with the University of Zagreb and a full member of the Croatian Academy of Technical Sciences. He led a series of research projects financed by industry and different government agencies. He published more than 100 articles in scientific journals and conference proceedings in the area of modeling and simulation, guidance and control, defense systems and technologies, interactive simulation systems, virtual reality, affective computing, and cyber psychology. He was also a Codirector of a series of conferences in the framework of NATO Science for Peace programs.

SINI A POPOVI received the B.Sc. degree in mathematics, and the M.Sc. and Ph.D. degrees in computer science, in 2001, 2006, and 2011, respectively. He is currently an Assistant Professor with the Faculty of Electrical Engineering and Computing, University of Zagreb. He has mostly worked on applications of virtual reality and/or affective computing methods in psychotherapy of stress-related disorders and assessment of human resilience to stress. His research interests include affective computing, interactive simulation, virtual reality, and their intersections with psychology and neuroscience.

MARKO ARLIJA received the B.Sc. and M.Sc. degrees in electrical engineering and information technology, in 2014 and 2016, respectively. He is currently pursuing the Ph.D. degree in computer science. He is currently a Research and Teaching Assistant with the Faculty of Electrical Engineering and Computing, University of Zagreb. He has worked most extensively in the area of physiology-based stress resilience prediction. His research interests include system theory, physiological signal processing, affective computing and machine learning applications, and lying at the intersection with psychophysiology and neuroscience.

IGOR MIJI received the B.S. and M.S. degrees in electrical engineering from the Faculty of Electrical Engineering and Computing, University of Zagreb, in 2012 and 2014, respectively, where he is currently pursuing the Ph.D. degree in computing. He is currently a Data Scientist in the field of natural language processing with RealNetworks. His research interests include speech processing, affective computing, and natural language processing.

MIRKO KOKOT received the B.Sc. and M.Sc. degrees in electrical engineering and information technology, in 2015 and 2017, respectively. He is currently pursuing the Ph.D. degree in electrical engineering with the Faculty of Electrical Engineering and Computing, University of Zagreb. He is a Research and Development Engineer with RoMb Technologies d.o.o. Previously, he was a Researcher in the field of physiological signal processing and affective computing. His research interests include control theory, robotics, and multiagent systems. He was a recipient of the 2016 Rector's Award of University of Zagreb for his work in applying affective computing in robotics.

IVAN KESED I received the B.Sc. and M.Sc. degrees in information and communication technology, in 2014 and 2016, respectively. He is currently pursuing the Ph.D. degree in computing. He is a Research and Teaching Assistant with the Faculty of Electrical Engineering and Computing, University of Zagreb. He has worked most extensively on brain-imaging analysis in cognitive tasks and resilience to stress prediction. His research interest includes signal and image processing.

GARY STRANGMAN received the B.S. degree in mathematics and chemistry, in 1991, and the M.S. and Ph.D. degrees in cognitive science, in 1994 and 1998, respectively. He is currently an Associate Professor with the Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School. He has worked extensively with NASA to develop cognitive and neurophysiology tools to support the monitoring, maintenance, and enhancement of behavioral health in spaceflight. His research interests include neurophysiology and human performance, particularly in extreme environments.

VLADIMIR IVKOVI received the B.A. degree in psychology, in 1999, the M.Sc. degree in biology, in 2005, the M.Sc. degree in space studies, in 2006, and the Ph.D. degree in neuromotor control, cognitive science, in 2012. He is currently an Instructor with the Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School. He has worked extensively with NASA, Boston Fire Department, and other agencies on development of countermeasures for mitigation of operational and occupational performance decrements and health risks. His research interests include ambulatory brain and physiology monitoring for assessment of neurophysiologic and neurobehavioral disorders elicited by exposure to extreme operational environments or activities, including spaceflight, firefighting, and emergency response.

QUAN ZHANG (M'06) was born in Luoyang, China, in 1972. He received the B.S. and Ph.D. degrees in biomedical engineering from Xi'an Jiaotong University, China, in 1993 and 1997, respectively, and the postdoctoral fellowships from the University of Pennsylvania and from the Massachusetts General Hospital (MGH) and Harvard Medical School (HMS). He is currently the Director of the NSG Biomedical Engineering Laboratory, MGH, and an Instructor in psychology with HMS. His research interests include the development and applications of novel wearable functional brain monitoring, neuroimaging, and unobtrusive blood pressure monitoring technologies.