Unraveling the Physiological Correlates of Mental Workload Variations in Tracking and Collision Prediction Tasks

Alka Rachel John, Avinash K. Singh, *Member, IEEE*, Tien-Thong Nguyen Do, Ami Eidels, Eugene Nalivaiko, Alireza Mazloumi Gavgani, Scott Brown, Murray Bennett, Sara Lal, Ann M. Simpson, Sylvia M. Gustin, Kay Double, Frederick Rohan Walker, Sabina Kleitman, John Morley, and Chin-Teng Lin, *Fellow, IEEE*

***Abstract*— Modern work environments have extensive interactions with technology and greater cognitive complexity of the tasks, which results in human operators experiencing increased mental workload. Air traffic control operators routinely work in such complex environments, and we designed tracking and collision prediction tasks to emulate their elementary tasks. The physiological response to the workload variations in these tasks was elucidated to untangle the impact of workload variations experienced by operators. Electroencephalogram (EEG), eye activity, and heart rate variability (HRV) data were recorded from 24 participants performing tracking and collision prediction tasks with three levels of difficulty. Our findings indicate that variations in task load in both these tasks are sensitively reflected in EEG, eye activity and HRV data. Multiple regression results also show that operators' performance in both tasks can be predicted using the corresponding EEG, eye activity and HRV data. The results also demonstrate that the brain dynamics during each of these tasks can be estimated from the corresponding eye activity, HRV and performance data. Furthermore, the markedly distinct neurometrics of workload variations in the tracking and collision prediction tasks indicate that neurometrics can provide insights on the type of mental workload. These findings have applicability to the design of future mental workload adaptive systems that integrate neurometrics in deciding not just "when" but also "what" to adapt. Our study provides compelling evidence in the viability of developing intelligent closed-loop mental workload adaptive systems that ensure efficiency and safety in complex work environments.**

***Index Terms*— Mental workload, EEG, pupil size, blink rate, RMSSD.**

Manuscript received August 30, 2021; revised January 14, 2022; accepted February 27, 2022. Date of publication March 8, 2022; date of current version March 28, 2022. This work was supported in part by the Australian Research Council (ARC)under Discovery Grant DP210101093 and Discovery Grant DP220100803, in part by the Australia Defence Innovation Hub under Contract P18-650825, in part by the U.S. Office of Naval Research Global through Cooperative Agreement under Grant ONRG-NICOP-N62909-19-1-2058, in part by AFOSR-DST Australian Autonomy Initiative Agreement under Grant ID10134, and in part by NSW Defence Innovation Network and NSW State Government of Australia under Grant DINPP2019 S1-03/09 and Grant PP21-22.03.02. *(Corresponding author: Chin-Teng Lin.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the University of Technology Sydney Human Research Ethics Expedited Review Committee under Application No. ETH19-4197.

Alka Rachel John, Avinash K. Singh, Tien-Thong Nguyen Do, and Chin-Teng Lin are with the Australian Artificial Intelligence Institute, FEIT, University of Technology Sydney, Ultimo, NSW 2007, Australia (e-mail: chin-teng.lin@uts.edu.au).

Ami Eidels, Scott Brown, and Murray Bennett are with the School of Psychology, University of Newcastle, Callaghan, NSW 2308, Australia.

Eugene Nalivaiko, Alireza Mazloumi Gavgani, and Frederick Rohan Walker are with the School of Biomedical Sciences and Pharmacy, and the Centre for Advanced Training Systems, University of Newcastle, Callaghan, NSW 2308, Australia.

Sara Lal and Ann M. Simpson are with the School of Life Sciences, University of Technology Sydney, Ultimo, NSW 2007, Australia.

Sylvia M. Gustin is with the School of Psychology, University of New South Wales, Sydney, NSW 2052, Australia, and also with the Centre for Pain IMPACT, Neuroscience Research Australia, Randwick, NSW 2031, Australia.

Kay Double is with the Brain and Mind Centre and Discipline of Pharmacology, School of Medical Sciences, The University of Sydney, Sydney, NSW 2006, Australia.

Sabina Kleitman is with the Faculty of Science, The University of Sydney, Sydney, NSW 2006, Australia.

John Morley is with the School of Medicine, Western Sydney University, Penrith, NSW 2751, Australia.

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSRE.2022.3157446, provided by the authors.

Digital Object Identifier 10.1109/TNSRE.2022.3157446

I. Introduction

MENTAL workload is one of the most crucial factors that affect the efficiency of human operators as they function in complex interactive work environments. Wickens and Tsang [1] defined mental workload as the dynamic relationship between the cognitive resources demanded by a task and the capability of the operator to afford those resources.

The theory of limited cognitive resources states that exposure to demanding task conditions impairs performance due to resource depletion [2] or compromised access to resources [3]. As mental workload has a negative influence on the performance of the operator, it results in human error commission [4], compromising system efficiency and safety [5]. Mental workload must be maintained at an optimal level, avoiding both underload and overload [6] as the performance is known to fall at both overload and underload conditions [6], [7].

Predicting an operator's mental workload and thereby adapting the system behaviour by modifying task allocation can avoid the loss of situational awareness, maintaining high performance. Accurate and reliable measurement of the mental workload of an operator is crucial, especially in a safety-critical work environment, by providing better work environments and human-machine interactions [8], [9].

Researchers have relied on multiple strategies, such as self-assessment, performance measures and physiological metrics, to assess mental workload; however, each of these methods has its benefits and drawbacks [10]. Over the years, physiological metrics have been used to assess workload [11], [12] as it offers high sensitivity, diagnostic ability and is mostly non-intrusive [13], providing an accurate and real-time assessment of the operator's workload. The use of physiological data such as neurophysiological signals can assess mental workload online without influencing the task as there is no explicit output [14], [15]. Neurophysiological measures can also assess the changes in the mental state that are not merely discernible in overt task performance [1], [15], [16].

Neurophysiological measures, such as the electroencephalogram (EEG) signal, has been widely employed to estimate mental workload as the effects of task demand are clearly visible in EEG rhythm variations [14], [17]–[19]. Researchers have also used EEG to predict performance degradation from workload variations reliably [21], [22] and noted that it is correlated with an increase in the frontal theta power and a change in parietal alpha power, which relates to cognitive and memory performance [14], [21]–[23]. However, EEG features of the mental workload are found to be task-dependent, therefore, adding other modalities like eye activity data and heart rate data can help achieve far superior outcomes [24].

This paper investigated whether the multimodal physiological metrics of mental workload can provide more information about the task contributing to the workload. We designed tracking and collision prediction tasks to elucidate the physiological effects of workload variations in these tasks. The tasks were inspired by the real-world tasks that air traffic control (ATC) operators routinely perform to ensure a safe and efficient air traffic flow. Even though several factors influence the complexity of ATC tasks [25], [26], such as environmental, display, traffic and organisational aspects, the main functions of ATC operators are tracking and collision prediction.

The tracking task was designed based on a widely employed and extensively researched paradigm of multiple object tracking [27], [28]. This tracking task was chosen as it emulates the real-world ATC job of keeping track of aircraft, and the design of the collision prediction task was inspired by the conflict detection task of ATC.

The experiment was fashioned as a classical cognitive paradigm with manipulation of workload (low, medium, high) and repeated stimuli to study whether physiological data such as EEG, eye activity and HRV can reliably assess the mental workload of the operator while they perform these basic tracking and collision prediction tasks. We formulated the following four research hypotheses for our study:

- H1 The three distinct levels of workload defined in both tracking and collision prediction tasks can yield significant performance degradation with the increasing levels of workload.
- H2 Workload variation in tracking and collision prediction tasks can be reliably assessed using EEG, eye activity and HRV metrics.
- H3 The performance in tracking and collision prediction tasks can be predicted based on the measured physiological signals.
- H4 Physiological response to the workload variations in the tracking and collision prediction tasks will be distinct across tasks.

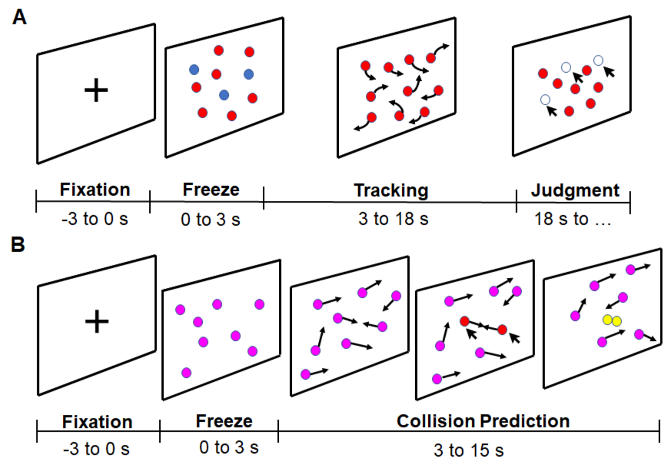

Fig. 1. The experimental design of the tasks. (A) the experimental design of the tracking task and (B) shows the design of the collision prediction task. The number of dots shown in these diagrams is just for representation purposes.

II. Materials and Methods

A. Participants

Twenty-four participants (age 25 ± 5, 17 males and seven females, all right-handed), recruited at the University of Technology Sydney, participated in this experiment after giving written informed consent. Participants did not have prior knowledge of the experimental scenario. All the participants had normal or corrected vision and no history of any psychological disorder that might affect the results. The experimental protocol was approved by the University of Technology Sydney Human Research Ethics Expedited Review Committee (ETH19-4197).

EEG data was collected using the SynAmps2 Express system (Compumedics Ltd., VIC, Australia) with 64 Ag/AgCl sensors. Eye activity data was collected using Pupil Labs (Pupil Core, Berlin, Germany). This wearable eye-tracking headset has three cameras, two of which record the eyes activity at a 200 Hz sampling rate, and the other one records the participant's field of view at a 30 Hz sampling rate [29]. The Blood Volume Pulse (BVP) data was recorded using infrared plethysmography-based Empatica E4 (Empatica Srl, Milano, Italy). The real-time synchronisation of events from the task scenario to the EEG, eye activity and BVP data was achieved by the Lab Streaming Layer [30].

B. Experimental Procedure

Our experimental design included two tasks – multiple objects tracking task [31] and collision prediction task. As shown in Fig. 1(A), in the tracking task, during the initial

TABLE I
WORKLOAD MANIPULATIONS IN THE TRACKING AND COLLISION PREDICTION TASKS

Task	Workload Level	Workload Manipulation	
		Tracking Dots	Total Dots
Tracking Task	Low	1	10
	Medium	3	12
	High	5	15
		Workload Manipulation	
Collision Prediction Task	Low	6	
	Medium	12	
	High	18	

3 seconds, participants look at a fixation cross on the screen followed by a freeze phase, where the dots, some of which are blue, and the rest are red, remain stationary. The blue dots are the dots that should be tracked (hence, 'targets'). After three seconds of freeze, the blue targets also turn red so that they are no longer distinctive from the other dots and all the dots start moving. Each dot has a diameter of 14 pixels, and they move randomly in the display area at a frame rate of 15 frames/second. The participant is asked to keep track of the targets (initially blue dots) for 15 seconds. After this time window, all dots stop moving, and the participants should indicate the target dots by clicking on the dots that they have kept track of. The workload levels in this tracking task are manipulated by varying the number of blue dots and the total number of dots, as shown in Table I.

As shown in Fig. 1(B), in the collision prediction task, there is a fixation cross on the screen for three seconds. Then there is a three-second-long freeze phase where the dots remain stationary, after which all the dots start moving. All dots are of the same colour (pink), unlike the tracking task. The participant is required to predict the trajectory of the dots and identify which pair of dots would collide. Dots move in a predefined uniform direction at a frame rate of 15 frames/second, and we have manipulated the trajectory of the dots such that there will be only one collision in each trial. The participants were asked to identify the pair of dots that would collide and click on both dots before the collision, which usually occurs in the last 3 seconds of the trial. In order to prevent random guesses, the number of dots the participants can select is limited to two, and once the participant clicks on the dot, it changes from pink to red colour. The levels of workload were manipulated by varying the number of dots as shown in Table I. A 15-inch monitor with 1920 x 1080 resolution was used to display both these tasks. We carried out the experiment in a light-controlled room, which provided a good balance of luminance in the experimental environment and the display screen, avoiding any direct or reflected glare. Furthermore, the tracking and collision prediction tasks were designed with a black background to avoid any eye fatigue effects.

Each participant had to perform 108 trials of each task with 36 trials of each workload level. The entire experiment was divided into four blocks, and each block had 27 trials of the tracking task and 27 trials of the collision prediction task. The type of workload condition in the trials was randomised within a block to avoid habituation or expectation effects.

After each block, the participants were advised to rest for five minutes before proceeding to the next block by pressing the spacebar key. Also, within a block, after completing each trial, participants move to the subsequent trial by pressing the spacebar key. The participants were advised to self-pace and rest before proceeding to the next trial to avoid fatigue. Furthermore, to avoid measuring any fatigue effects, the EEG, eye activity and HRV data from each trial were normalised by considering the fixation period at the beginning of every trial as the baseline.

All participants were trained in a training session that lasted approximately ten minutes, where they performed six trials of each task to familiarise themselves with the tasks and develop strategies for successfully executing the tasks. The participants were asked to continue the training until they felt comfortable with the tasks. After the training, all participants performed the tasks for $\tilde{1}$.5 hours, during which EEG, eye activity and HRV data were collected.

C. Data Analysis

1) Behavioural and Performance Data Analysis: For the tracking task, performance was evaluated by examining tracking accuracy. The tracking accuracy for each trial was defined as the ratio of the number of correctly tracked dots to the total number of dots to track.

$$\text{Tracking Accuracy} = \frac{\text{Number of Correctly Tracked Dots}}{\text{Total Number of Dots to Track}} \tag{1}$$

The performance in the collision prediction trials was determined using the time before collision and collision miss proportion rate. The time before collision is the time period between when the participant clicks on either one of the colliding dots and when the collision happens. The collision miss proportion rate for a particular workload level of the collision prediction task is the ratio of the number of collision prediction misses to the total number of collisions in that specific workload level. A collision miss was considered to happen when the participant could not identify which pair of dots would collide and, hence, did not click on either of the dots before the collision.

$$\text{Collision Prediction Miss Proportion Rate} = \frac{\text{Number of Missed Collision Predictions}}{\text{Total Number of Collisions}} \tag{2}$$

2) EEG Preprocessing: EEG data were preprocessed using EEGLAB v2020.0 toolbox [32] in MATLAB R2019a (The Mathworks, Inc., Natick, MA, USA) and adapted from [33]. EEG data were down-sampled to 250 Hz, and a band-pass filter of 2–45 Hz was applied. Channels with three seconds or more flat line were removed using the clean_flatline function. Noisy channels were identified and removed using the clean_channels function in EEGLAB. On average, 3±1 channels were removed, and these channels were restored by interpolating the data from neighbouring channels using the spherical spline method from the EEGLAB toolbox. Continuous artifactual regions were removed using the EEGLAB function, pop_rejcont. Then window cleaning was performed using the clean_windows function in EEGLAB.

After these artifact removal steps, two EEG datasets were extracted, one comprising tracking trials and one with the collision prediction trials. Each participant had 30±2 high, 31±1 medium and 29±1 low workload tracking trials, and 32±2 high, 32±2 medium and 30±1 low workload collision prediction trials.

The tracking epochs were 21 seconds long and included the three seconds of fixation period followed by the three seconds of freeze, after which the tracking task was commenced. The collision prediction task epochs were 15 seconds in length and included the initial three seconds of fixation period followed by the three seconds of freeze and then the collision prediction task. Both tracking and collision prediction datasets were decomposed using Independent Component Analysis (ICA), performed using EEGLAB's runica algorithm [32]. Finally, we employed ICLabel [34], an automatic IC classifier to identify components related to brain, heart, line noise, eye, muscle, channel noise and other activities. This tool was adopted to generate class labels for each component, and all the components with labels other than brain activity were rejected.

a) IC Clustering: EEGLAB STUDY structure [35] was used to manage and process data recorded from multiple participants as it provides component clustering to cluster similar independent components across participants and allows statistical comparisons of component activities for different workload conditions. Clustering functions were used to examine the contributions of frontal, parietal and occipital clusters of independent components (ICs) to the workload dynamics. Frontal and parietal brain regions have been reported to reflect the changes in workload [11], [17], [19], [36]–[38], and as both our tasks also manipulate the visual load, we mainly focused on the frontal, parietal and occipital clusters of brain activity.

A Study was created for each task, and each Study had one group (with 24 participants) with three conditions corresponding to the three levels of workload. Since the dataset of each participant was recorded in a single session, the resulting independent component maps were the same across all three conditions for each participant. For each participant, only those ICs that had a residual variance (RV) less than 15% and inside the brain volume were chosen, which was achieved using Fieldtrip extension [39]. The k-means clustering algorithm [40] was used to cluster independent components across all participants to clusters based on two equally weighted (weight=1) criteria: (1) scalp maps and (2) their equivalent dipole model locations, which was performed using DIPFIT routines [41] in EEGLAB. Talairach coordinates [42] of the fitted dipole sources of these clusters were identified to select frontal, parietal and occipital clusters.

The grand-mean IC event-related spectral power changes (ERSPs) for each condition was subsequently calculated for each cluster. Fixation phase in each tracking and collision prediction epoch was taken as the baseline to see the changes in power spectra during the task. ERSPs for frontal, parietal and occipital clusters for tracking and collision prediction tasks were examined. To compare the ERSP of different workload conditions, permutation-based statistics, implemented in EEGLAB, was used with Bonferroni correction and significance level set to $p = .05$.

Also, for the frontal, parietal and occipital clusters, each ICs' spectral powers were calculated using EEGLAB's spectopo function, which uses Welch's periodogram method [43] on each 2-s segment using a Hamming window with 25% overlap for a range of frequencies from 2 to 45 Hz. For each IC, the power spectral density (PSD) at different frequency bands were examined to identify the correlates of mental workload, and the results are available in the Supplementary Material.

3) Eye Activity Data: Pupil Core software, Pupil Capture, provides the pupil size for the left and right eye separately along with the associated confidence value, representing the quality of the detection result. All data points where the confidence of the pupil size was less than 0.8 were removed from the data. The pupil size data was low pass filtered (using a minimum order finite impulse response filter) at 4 Hz [44]. The raw pupil size data was normalised using the baseline data (defined as the three seconds of fixation period in each tracking and collision prediction epoch). The blinks during each trial was also extracted from the pupil size measurement when the pupil size and confidence of the measurement, reported by the Pupil Capture software, suddenly dropped to zero.

4) Heart Rate Variability: Inter-beat-interval (IBI) time series was computed from the Blood Volume Pulse (BVP) data of each tracking and collision prediction trial. Root Mean Square of the Successive Differences (RMSSD) was computed by detecting peaks of the BVP using PeakUtils Python package [45] and calculating the lengths of the intervals between adjacent beats.

$$RMSSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(IBI_{i-1} - IBI_i)^2} \tag{3}$$

RMSSD data was also normalised by considering the fixation period in each tracking and collision prediction epoch as the baseline.

5) Statistical Analysis: Statistical analyses were carried out using the SPSS (IBM SPSS 26.0; Chicago, IL, U.S.A.) statistical tool. In order to investigate the differences in the performance, EEG, eye activity and HRV parameters across participants in the three workload levels of tracking and collision prediction tasks, repeated-measures analysis of variance (ANOVA) was conducted with workload level (low, medium or high) as the within-subjects factor. Mauchly's test was implemented to test for sphericity. We performed Greenhouse-Geisser correction if sphericity was not satisfied ($p < .05$). If the main effect of the ANOVA was significant, post-hoc comparisons were made to determine the significance of pairwise comparisons, using Bonferroni correction.

Principal Component Factor Analysis procedure was carried out to select factors and derive EEG and eye metrics for tracking and collision prediction tasks. Kaiser-Meyer-Olkin (KMO) test and Bartlett's test was carried out to ensure the suitability of the data for factor analysis. The varimax method was used for factor rotation. Only the factors satisfying

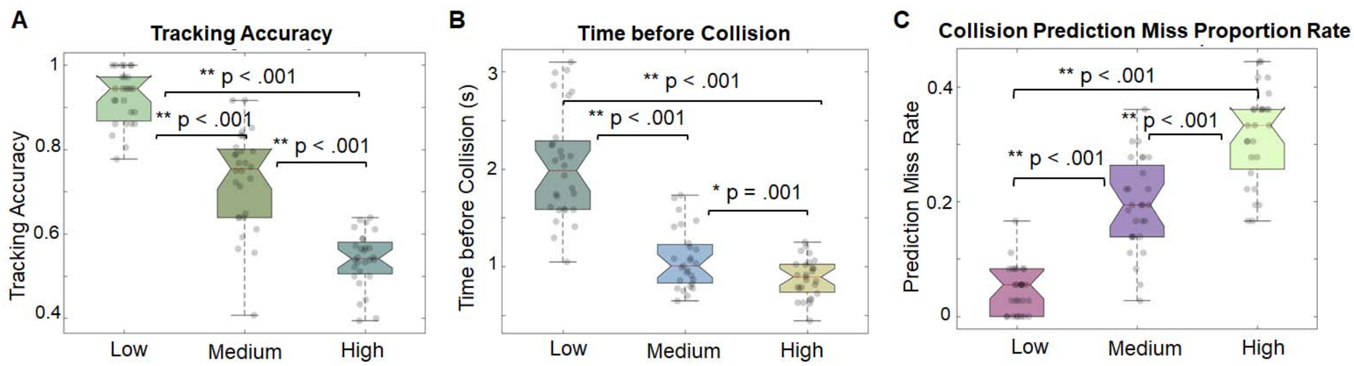

Fig. 2. (A) shows the tracking accuracy of all the participants in the tracking task for the three levels of workload. (B) and (C) shows the performance of all participants in the collision prediction task for the three levels of workload. (B) shows the mean time before collision for all the participants in the low, medium, and high workload conditions. (C) shows the collision prediction miss proportion rate for the three levels of workload.

Kaiser–Guttman criterion were included in determining EEG and eye metrics in both tasks.

Finally, multiple linear regression was performed to relate EEG, eye activity and HRV metrics to the performance in the tracking and collision prediction tasks. EEG power, eye activity and HRV metrics were all entered as predictors using the enter method, and the performance in the task was the dependent variable. Multiple linear regression was also performed to relate task performance, eye activity and HRV metrics to the observed brain dynamics during the tracking and collision prediction tasks. Performance, eye activity and HRV metrics were all entered as predictors using the enter method, and the brain dynamics in the task was the dependent variable.

III. Results

A. Behavioural and Performance Measures

In the tracking task, tracking accuracy decreased significantly with increasing levels of workload, as shown in Fig. 2(A). A repeated-measures ANOVA showed that tracking accuracy differed significantly between workload conditions [$F(2, 48) = 239.910$, $p < .001$, $\eta_p^2 = .899$].

For the collision prediction task, the time before collision and collision prediction miss proportion rate was considered. The time before collision decreased with increasing workload, as shown in Fig. 2(B). A repeated-measures ANOVA was conducted to study the effect of workload variations on time before collision, and the results showed that time before collision varied significantly between workload conditions [$F(1.497, 40.406) = 132.688$, $p < .001$, $\eta_p^2 = .831$]. The collision prediction miss proportion rate increased with increasing levels of workload, as shown in Fig. 2(C). Repeated-measures ANOVA showed that the collision prediction miss proportion varied significantly between workload conditions [$F(1.593, 43.009) = 116.338$, $p < .001$, $\eta_p^2 = .812$].

B. EEG Results

1) Independent Source Clusters: The frontal, parietal and occipital clusters were selected based on the location of fitted dipole sources [41]. For the tracking task (refer Fig. 3(A1)), the Talairach coordinate of the frontal, parietal and occipital clusters centroid were at (−1, 41, 27), (4, −51, 39) and (30, −70, 15) respectively (refer Fig. 3(B1)).

For the collision prediction task (see Fig. 4(A1), 4(B1) and 4(C1)), the Talairach coordinate of the frontal, parietal and occipital clusters centroid were at (−10, 17, 46), (5, −47, 47) and (−3, −69, 20) respectively.

2) ERSP Changes With Mental Workload: Fig. 3(A2) and 3(B2) illustrates frontal and occipital clusters' ERSP changes for three workload conditions: low, medium and high during the tracking task. Statistical analysis on ERSP changes of the frontal cluster are shown in Fig. 3(A3)). It revealed a significant increase in theta power from the low to the high level ($p < .05$) and a significant increase in theta power at the frontal cluster during the high workload condition compared to the medium workload condition. The frontal theta power during the medium workload condition was significantly greater than the low workload condition.

However, no significant spectral power variations were observed at the parietal cluster. Fig. 3(B2) shows the ERSP changes at the occipital cluster. Fig. 3(B3) reveals the results of statistical analysis on the ERSP changes at the occipital cluster. It was revealed that there was a significant decrease in alpha power from the low to the high level ($p < .05$) and a significant decrease in alpha power at the occipital cluster during the high workload condition compared to the medium workload condition. The occipital alpha power during the medium workload condition was significantly less than the low workload condition.

Fig. 4(A2), 4(B2) and 4(C2) illustrates the frontal, parietal and occipital clusters' ERSP changes for three workload conditions in the collision prediction task. Statistical analysis on ERSP changes of the frontal cluster showed a significant increase in theta power during the high workload condition as compared to the low workload condition (Fig. 4(A3)). The frontal theta power during the high workload condition was also significantly greater than that of the medium workload. Further, there was a significant increase in the frontal theta power during medium workload compared to the low workload condition in the collision prediction task.

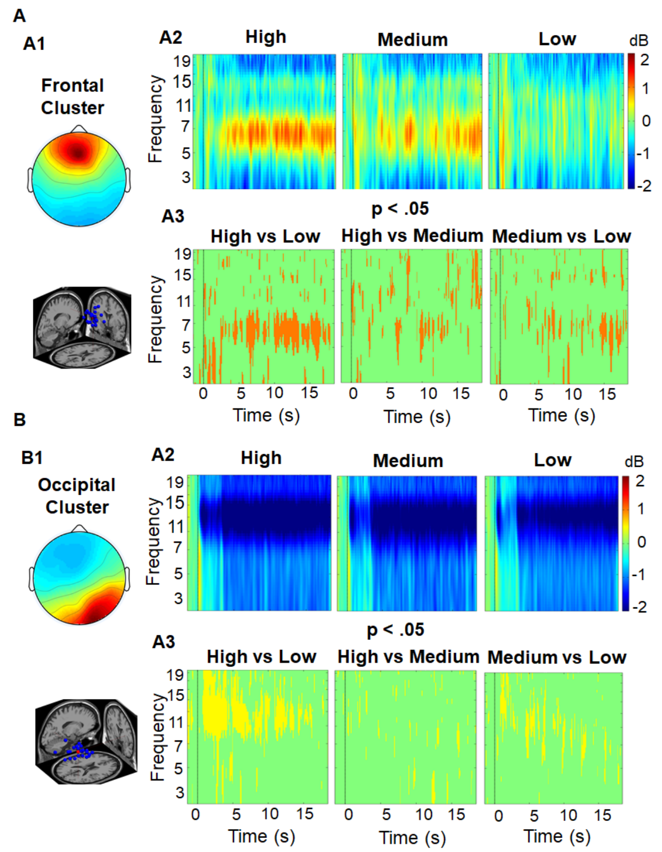

Fig. 3. Scalp map, dipole locations and ERSP changes at the Frontal and Occipital clusters selected in the tracking task. (A1) spatial scalp map and dipole source locations of the frontal cluster. (A2) shows the ERSP changes at the frontal cluster during high, medium and low workload conditions. (A3) shows the statistically significant difference ($p < .05$) between high and low workload conditions, high and medium workload conditions, and medium and low workload conditions. (B1) spatial scalp map and dipole source locations of the occipital cluster. (B2) shows the ERSP changes at the occipital cluster during high, medium, and low workload conditions. (B3) shows the statistically significant difference ($p < .05$) between high and low workload conditions, high and medium workload conditions and medium and low workload conditions.

The statistical analysis on the ERSP changes at the parietal cluster (Fig. 4(B3)) revealed a significant increase in the theta power in high workload as compared to low workload condition ($p < .05$) and a significant decrease in the alpha power ($p < .05$). There was a significant increase in the theta power and a significant decrease in the alpha power at the parietal cluster during the high workload condition compared to the medium workload condition. In the medium workload condition, while the parietal theta power was significantly higher, the parietal alpha power was significantly lower than the low workload condition. The ERSP changes at the occipital cluster (Fig. 4(C3)) revealed a significant increase in the delta and theta power in the high workload as compared to the low workload condition ($p < .05$). There was also a significant increase in the delta and theta power at the occipital cluster during the high workload condition compared to the medium workload condition. In the medium workload condition, the occipital delta and theta power were significantly higher than in the low workload.

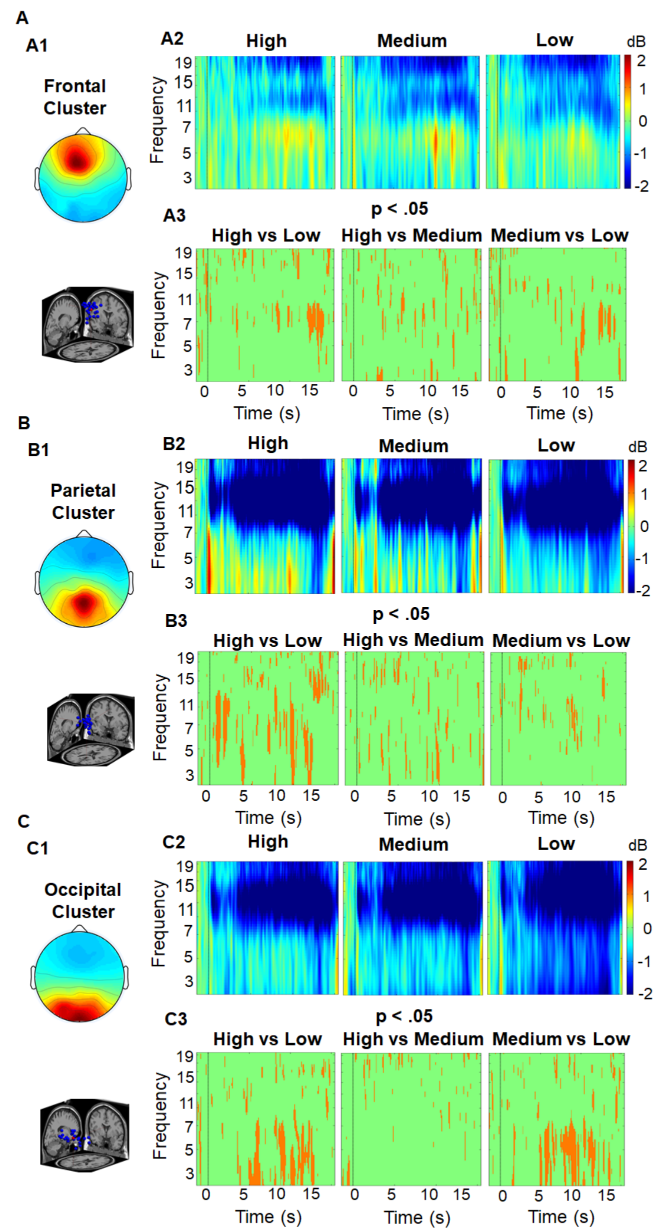

Fig. 4. Scalp map, dipole source locations and ERSP changes at the Frontal, Parietal and Occipital clusters selected in the collision prediction task. (A1) spatial scalp map and dipole source locations of the frontal cluster. (A2) shows the ERSP changes at the frontal cluster during high, medium, and low workload conditions. (A3) shows the statistically significant difference ($p < .05$) between the high and low workload conditions, high and medium workload conditions, and medium and low workload conditions. (B1) spatial scalp map and dipole source locations of the parietal cluster. (B2) shows the ERSP changes at the parietal cluster during high, medium, and low workload conditions. (B3) shows the statistically significant difference ($p < .05$) between high and low workload conditions, high and medium workload conditions, medium and low workload conditions. (C1) spatial scalp map and dipole source locations of the occipital cluster. (C2) shows the ERSP changes at the occipital cluster during high, medium, and low workload conditions. (C3) shows the statistically significant difference ($p < .05$) between the high and low workload conditions, high and medium workload conditions, and medium and low workload conditions.

C. Eye Activity Changes With Mental Workload

As shown in Fig. 5(A), pupil size increased with the increasing workload for both tracking and collision prediction tasks.

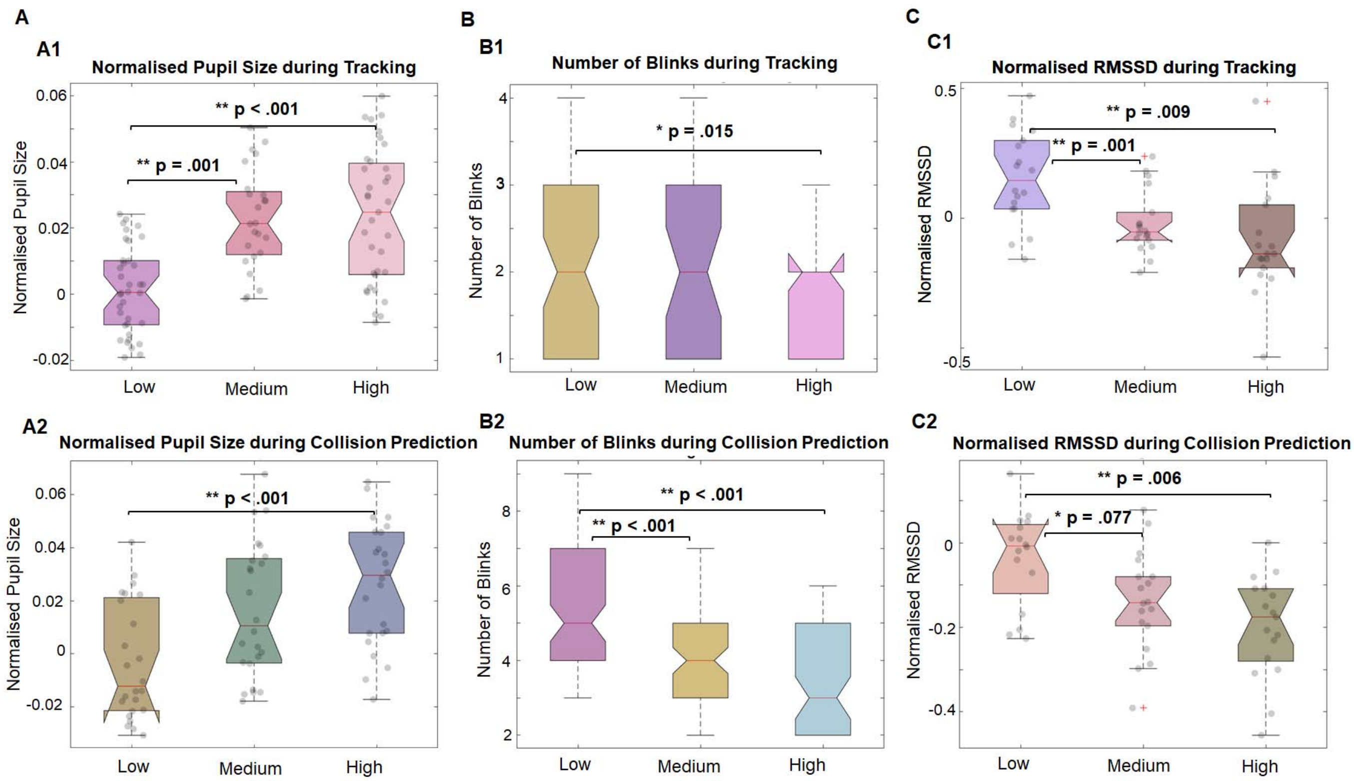

Fig. 5. (A) shows the normalized pupil size of all the participants shows a positive trend with the increasing workload. (A1) Normalised pupil size in the three workload conditions of the tracking task. (A2) Normalised pupil size during low, medium, and high workload conditions for the collision prediction task. (B) shows the negative trend in the number of blinks with the increasing workload. (B1) Number of blinks during different workload conditions of the tracking task. (B2) Number of blinks during the collision prediction task decreases with increasing level of workload. (C) shows the declining trend in the normalized RMSSD of all the participants with the increasing workload. (C1) Normalised RMSSD all the participants in the low, medium, and high workload conditions of the tracking task. (C2) Normalised RMSSD during collision prediction task for the three levels of workload.

For the tracking task, there was a significant change in the pupil size for different workload conditions as shown by repeated-measures ANOVA [$F(2, 38) = 13.205$, $p < .001$, $\eta_p^2 = .410$]. The results of repeated measures ANOVA shows that in the collision prediction task, there was a significant change in the pupil size for different workload conditions [$F(2, 46) = 9.276$, $p < .001$, $\eta_p^2 = .287$].

The number of blinks during tracking and collision prediction tasks decreased with the increasing workload, as shown in Fig. 5(B). Repeated-measure ANOVA was conducted to study the effect of workload variations on the number of blinks, which revealed significant variations in the number of blinks for different workload conditions during the tracking task [$F(2, 46) = 3.624$, $p = .035$, $\eta_p^2 = .136$]. The effect of workload on the number of blinks in the collision prediction task was analysed using repeated-measure ANOVA. It showed a significant variation in the number of blinks [$F(2, 46) = 18.586$, $p < .001$, $\eta_p^2 = .447$].

D. Heart Rate Variability (RMSSD) Changes With Mental Workload

Fig. 5(C) shows the RMSSD variation for different workload conditions in the tracking and collision prediction task. For the tracking task, there was a significant change in the RMSSD for the different workload conditions, as shown by the repeated-measures ANOVA [$F(2, 34) = 10.171$, $p < .001$, $\eta_p^2 = .374$]. Results from repeated-measures ANOVA also shows that in the collision prediction task, there was a significant change in the RMSSD for different workload conditions [$F(2, 44) = 4.279$, $p = .022$, $\eta_p^2 = .201$].

E. Performance Can Be Predicted From Physiological Data - Multiple Regression Results

Performance measure in the collision prediction task was derived using Principal Component Factor Analysis procedure with the time before collision and collision prediction miss proportion rate as variables. The eye activity metric for tracking and collision prediction tasks were derived using factor analysis with pupil size and blink rate as variables. The EEG metric was derived through factor analysis with frontal theta PSD and occipital alpha PSD as variables for the tracking task. The EEG metric was derived using factor analysis with frontal theta, parietal theta and alpha, occipital delta and theta PSD as variables for collision prediction. Only the factors with eigenvalue > 1 were included to derive the metrics.

Multiple regression was carried out to investigate whether EEG, eye activity and HRV metrics of workload could significantly predict the performance in the tracking task. The regression results indicated that the model explained 54.3% of the variance and that the model was a significant predictor of the tracking performance, $F(3, 67) = 26.543$, $p < .001$. While EEG ($B = .067$, $p = .001$) and eye activity ($B = -.089$, $p < .001$) contributed significantly to the model, HRV metric did not ($B = -.152$, $p = .125$). The final predictive model was:

$$\begin{aligned}&\text{Performance in tracking task}\\&\quad = 0.725 - 0.067 * \text{EEG metric}\\&\qquad - 0.089 * \text{Eye related metric} - 0.152 * \text{HRV metric} \qquad (4)\end{aligned}$$

In order to determine whether EEG, eye activity and HRV metrics could significantly predict the performance in collision prediction task, we conducted multiple regression analysis. The regression results indicated that the model explained 61.7% of the variance and that the model was a significant predictor of the performance in the collision prediction task, $F(3, 68) = 24.324$, $p < .001$. While eye activity ($B = -.276$, $p = .02$) and EEG metrics ($B = -.532$, $p < .001$) contributed significantly to the model, HRV metric did not ($B = .444$, $p = .443$). The final predictive model was:

$$\begin{aligned}&\text{Performance in Collision Prediction task}\\&\quad = 0.055 - 0.532 * \text{EEG metric} - 0.276\\&\qquad * \text{Eye related metric} + 0.444 * \text{HRV metric} \qquad (5)\end{aligned}$$

In order to determine whether brain dynamics in the tracking task can be predicted from the task performance, eye activity and HRV, multiple regression analysis was employed. The regression results indicated that the model explained 52.8% of the variance and that the model was a significant predictor of the brain dynamics, $F(3, 67) = 25.013$, $p < .001$. While performance in the task, evaluated by tracking accuracy ($B = -2.176$, $p = .001$) and eye activity ($B = .314$, $p = .015$) contributed significantly to the model, HRV metric did not ($B = -1.034$, $p = .067$). The final predictive model was:

$$\begin{aligned}&\text{Brain dynamics in tracking task}\\&\quad = 1.608 - 2.176 * \text{Performance}\\&\qquad + 0.314 * \text{Eye related metric} - 1.034 * \text{HRV metric} \qquad (6)\end{aligned}$$

Multiple regression analysis was also conducted to determine whether performance in the collision prediction task along with the eye activity and HRV metrics could significantly predict brain dynamics during the collision prediction task. The results of the regression indicated that the model explained 67% of the variance and that the model was a significant predictor of the brain dynamics during the collision prediction task, $F(3, 68) = 46.064$, $p < .001$. While eye activity ($B = .426$, $p < .001$) and task performance ($B = -.458$, $p < .001$) contributed significantly to the model, HRV metrics did not ($B = -.075$, $p = .889$). The final predictive model was:

$$\begin{aligned}&\text{Brain dynamics in Collision Prediction task}\\&\quad = -0.009\\&\qquad - 0.458 * \text{Performance} + 0.426 * \text{Eye related metric}\\&\qquad - 0.075 * \text{HRV metric} \qquad (7)\end{aligned}$$

IV. Discussion

In this study, we designed two simple tasks emulating the elementary ATC tasks: tracking and collision prediction tasks. Although both these tasks are inspired from the elementary tasks that ATC operators routinely perform in complex work environments, we considered them separately to untangle the differences in the physiological response to workload variations in these tasks.

In order to study the workload effects of increasing task load, the mental workload in both these tasks was manipulated by varying the number of dots. An increase in the mental workload results in several undesirable cognitive states such as inattention, mind wandering and effort withdrawal, reducing situational awareness [46], [47] which ultimately leads to poor performance and error commission [4], [48], [49]. Despite several studies [50]–[55] observing significant performance degradation with increasing workload levels, performance may not always reliably reflect workload variations as human operators can achieve the same performance experiencing different workload [56]. However, for both tasks in our experiment, the workload manipulation strategy for different workload levels was determined based on an initial pilot study conducted on eight participants. Only performance data was collected in these pilot studies, and the optimal number of dots for an effective workload manipulation for both tasks were determined based on performance. Further, the performance degradation with increasing workload levels in the tracking task was observed in a previous study [31]. A major limitation of this classical randomized workload experiment design is the absence of any self-assessment data, which would have provided more compelling evidence that the task load manipulations for both tracking and collision prediction tasks could elicit significant workload variations in the user.

It was observed that the performance in the tracking task, assessed by the tracking accuracy, degrades significantly with the increasing workload. Similarly, for the collision detection task, the time before collision decreased significantly, and collision prediction miss proportion also significantly increased with increasing levels of workload. Hence, we can confirm that the workload manipulation (by varying the number of dots) in both tracking and collision prediction tasks successfully elicited significant performance variations (H1).

In order to assess the mental workload, EEG, eye activity and BVP data were recorded while the participants performed the tasks. The component data was disentangled from the scalp EEG signal using independent component analysis. Significant correlations between mental workload and the spectral powers of frontal, parietal and occipital clusters were successfully elucidated.

The tracking task demands allocation of attentional resources to keep track of one, three or five tracking dots moving randomly among distractor dots. Researchers have suggested several theories regarding how human users can successfully track more than one moving target [57]–[64]. Working memory load is sensitive to increased allocation of attentional resources and is reflected by increases in frontal theta power [14], [65]. We observed an increase in the frontal theta power in the tracking task, which confirms that increased working memory load was experienced with increasing workload levels. Tracking dots moving among distractor dots also entails working memory mechanisms related to relevant item maintenance and increases in the memory load. This working memory mechanism is reflected by a decrease in the alpha power [21], [66]– [69]. The alpha power is also known to decrease with increased memory load [70]– [74] and task difficulty [23], [75]. Our findings also substantiate this working memory mechanism as the occipital alpha power decreases with increasing workload levels in the tracking task.

In the collision prediction task, anticipating the trajectory of dots and predicting whether dots would collide requires attention and internal concentration. Delta power is an indicator of attention or internal concentration in mental tasks, and it has been reported to increase with the increase in workload [23], [67], [76]. Our results demonstrate an increase in the delta power at the occipital sites, which validates an increased allocation of attentional resources with increasing levels of workload in the collision prediction task. Additionally, keeping a tab on the trajectory of six, 12 or 18 eight dots adds to the memory load in the participants. Several studies have shown that theta power is correlated with memory load [77], [78] and working memory capacity [64], [79], [80]. In the collision prediction task, our results reveal a significant increase in the theta power at the frontal, parietal and occipital clusters, confirming an increase in memory load with increasing workload levels. Furthermore, our results indicate that there is a decrease in parietal alpha power with increasing levels of workload. This observed alpha band desynchronisation with the increasing workload is related to relevant item maintenance in the working memory [21], [23], [69] and is known to decrease with increased memory load [70]– [74] and task difficulty [23], [75]. However, in the collision prediction task, the most significant decrease in the parietal alpha power was observed a few seconds before the collision. It might be related to the increase in the experienced time pressure [81] as the participants attempt to identify and click on the colliding pair of dots before the collision.

We also explored eye-related and HRV metrics during workload variations. Eye activity data was transformed to pupil size and blink rate information. Pupil size increased significantly with the increasing workload in both tracking and collision prediction tasks as the pupil dilates with increasing workload [82], [83]. The number of blinks also reduced considerably with the increasing workload in both tasks. Blink inhibition occurs in higher workload conditions [84] and so, the blink rate is inversely correlated with the attentional levels and workload experienced by the operator [17], [18], [67], [68]. RMSSD was negatively correlated with the mental workload in both tasks. This decrease in RMSSD with the increasing workload is widely reported in the literature [86], [87].

Our results show that EEG power spectra at the frontal, parietal and occipital areas, eye activity and HRV metrics can reliably and accurately assess the mental workload of the participants in both tasks. Hence, our second hypothesis (H2) is proved to be true for both tracking and collision prediction tasks. Relating to our third hypothesis (H3), the multiple regression results showed that the performance in the tracking and collision prediction tasks could be predicted from the EEG, eye-related and HRV metrics. However, only EEG and eye activity metrics contribute statistically to performance prediction in the tracking and collision prediction tasks. For the tracking task, the participants had a significant visual load in keeping track of a few dots moving among distractor dots which was reflected in the eye activity metric being a higher contributing predictor than the EEG metric. While for the collision prediction task, EEG metrics was a higher contributing predictor than eye activity metrics as participants require significant internal concentration to anticipate the trajectory of the dots and predict whether the dots would collide. The results also demonstrate that EEG and eye activity metrics explain more of the variance of performance in the collision prediction task than the tracking task. This might result from performance in the collision prediction task being derived from the time before collision and collision prediction miss proportion rate while performance in tracking task was represented by tracking accuracy alone.

Furthermore, our findings also demonstrate that the brain dynamics during each of these tasks can be estimated from the eye activity, HRV and performance during the tasks. However, only performance and eye activity metrics contribute statistically to the prediction of brain dynamics in the tracking and collision prediction tasks. Also, the variance of brain dynamics was better explained by performance and eye activity metrics in the collision prediction task than the tracking task. This might result from performance in the collision prediction task being derived from the time before collision and collision prediction miss proportion rate while performance in tracking task was represented by tracking accuracy alone. This estimation of the brain dynamics from the behaviour data will pave the way for future systems that can reliably estimate the brain activity from the behavioural response.

Even though EEG, eye activity and HRV measures were able to differentiate between low and high levels of workload sensitively, some of these measures could not accurately discern the medium workload from low/high workload conditions. There are two possible reasons for this incoherence reported in the literature: experiment design issue [88] or inter-individual differences [89]. In our experimental design, the medium workload condition might have required nearly comparable cognitive resources and hence, not qualifying for a significant variation from the low/high workload condition. However, our results showed a significant drop in the performance with increasing workload levels in both the tracking and collision prediction tasks.

Therefore, it is more plausible to reason that this incoherence might be due to the influence of inter-individual differences. It is well understood that the relationship between workload and task demand is not straightforward [90]. Sperandio [91] claims that the relationship can be better understood by investigating the strategies employed by human operators to manage their cognitive resources and workload. Therefore, the participants might reflect workload variations differently based on their cognitive resources and the strategies that they employ for performing the tasks.

Our results also indicate that even though eye activity and HRV metrics are sensitive to task load variations, they may not provide any valuable information on the task that causes the variations in workload. However, our results demonstrate that the EEG measures are not just sensitive to the workload variations but also the task type. The neurometrics correlated with the variations in the workload of tracking and collision prediction tasks are different, proving that our fourth hypothesis (H4) is true. Our results provide evidence that the use of EEG measures in a closed-loop adaptive system can not only aid the decision of "when" but also "what" form of automation to deploy to mitigate the workload variations in operators. Therefore, proper exploitation of the brain dynamics of an operator might help the adaptive automation to not just step in at the right time but also be cognitively empathetic with the operator, helping where it is needed, taking over the task that is currently overwhelming the operator. Hence, the results presented here contribute to the development of adaptive strategies essential for designing intelligent closed-loop mental workload adaptive systems.

V. Conclusion

The performance and efficiency of a system can be improved by maintaining the operator's workload in the optimal range. In order to elucidate the impact of task load variations that comprise the load variations in complex work environments, we separately designed two tasks emulating ATC tasks of tracking and collision prediction. EEG spectral power, eye and HRV correlates to mental workload variations for tracking and collision prediction tasks are successfully unravelled. Our results demonstrate that EEG, eye, and HRV metrics can provide a sensitive and reliable measure to predict the mental workload and performance of the operator. Furthermore, multiple regression results demonstrate that these physiological signals can predict task performance. The findings also reveal that the brain dynamics during each of these tasks can be estimated from the eye activity, HRV and performance during the tasks. The differences in neural response to increased workload in the tracking and collision prediction task indicate that these neural measures are sensitive to variations and type of mental workload. These differences demonstrate their potential utility in not just deciding "when" but also "what" to adapt, aiding the development of intelligent closed-loop mental workload aware systems. This investigation of physiological indices of workload variation in these basic tasks has applicability to the design of future adaptive systems that integrate neurometrics in deciding the form of automation to mitigate the variations in workload in complex work environments.

References

[1] C. Wickens and P. S. Tsang, "Workload," in *APA Handbook of Human Systems Integration*, D. A. Boehm-Davis, F. T. Durso, and J. D. Lee, Eds. Washington, DC, USA: American Psychological Association, 2015, pp. 277–292.

[2] H. P. A. Van Dongen, G. Belenky, and J. M. Krueger, "A local, bottom-up perspective on sleep deprivation and neurobehavioral performance," *Current Topics Med. Chem.*, vol. 11, no. 19, pp. 2414–2422, Sep. 2011.

[3] G. B. Pedraz, "Behavioural bases and functional dynamics of cognitive fatigue," Faculté des Sci. Psychologiques et de l'éducation, Université libre de Bruxelles, Brussels, Belgium, 2016.

[4] J. Reason, "Human error: Models and management," *BMJ*, vol. 320, no. 7237, pp. 768–770, Mar. 2000.

[5] B. Xie and G. Salvendy, "Prediction of mental workload in single and multiple tasks environments," *Int. J. Cognit. Ergonom.*, vol. 4, no. 3, pp. 213–242, Sep. 2000.

[6] P. A. Hancock, "A dynamic model of stress and sustained attention," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 31, no. 5, pp. 519–537, Oct. 1989.

[7] B. B. Van Acker, D. D. Parmentier, P. Vlerick, and J. Saldien, "Understanding mental workload: From a clarifying concept analysis toward an implementable framework," *Cognition, Technol. Work*, vol. 20, no. 3, pp. 351–365, Aug. 2018.

[8] E. A. Byrne and R. Parasuraman, "Psychophysiology and adaptive automation," *Biol. Psychol.*, vol. 42, no. 3, pp. 249–268, Feb. 1996.

[9] M. W. Scerbo, *The Efficacy of Psychophysiological Measures for Implementing Adaptive Technology*. Hampton, VA, USA: NASA Langley Research Center, 2001.

[10] R. D. O'Donnell and F. T. Eggemeier, "Workload assessment methodology," in *Handbook of Perception and Human Performance, Cognitive Processes and Performance*, vol. 2, K. R. Boff, L. Kaufman, and J. P. Thomas, Eds. New York, NY, USA: Wiley, 1986, pp. 42/41–42/49.

[11] G. Matthews, L. E. Reinerman-Jones, D. J. Barber, and J. Abich, "The psychometrics of mental workload: Multiple measures are sensitive but divergent," *Hum. Factors*, vol. 57, no. 1, pp. 125–143, 2014.

[12] R. L. Charles and J. Nixon, "Measuring mental workload using physiological measures: A systematic review," *Appl. Ergonom.*, vol. 74, pp. 221–232, Jan. 2019.

[13] R. Parasuraman and M. Rizzo, *Neuroergonomics: The Brain at Work*. London, U.K.: Oxford Univ. Press, 2008.

[14] A. Gevins and M. E. Smith, "Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style," *Cerebral Cortex*, vol. 10, no. 9, pp. 829–839, Sep. 2000.

[15] R. Parasuraman, "Neuroergonomic perspectives on human systems integration: Mental workload, vigilance, adaptive automation, and training," in *APA Handbook of Human Systems Integration*, D. A. Boehm-Davis, F. T. Durso, and J. D. Lee, Eds. Washington, DC, USA: American Psychological Association, 2015, pp. 163–176.

[16] B. Blankertz *et al.*, "The Berlin brain-computer interface: Progress beyond communication and control," *Frontiers Neurosci.*, vol. 10, p. 530, Nov. 2016.

[17] 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.

[18] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni, "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness," *Neurosci. Biobehav. Rev.*, vol. 44, pp. 58–75, Jul. 2014.

[19] G. Di Flumeri *et al.*, "EEG-based mental workload neurometric to evaluate the impact of different traffic and road conditions in real driving settings," *Frontiers Hum. Neurosci.*, vol. 12, p. 509, Dec. 2018.

[20] M. Matousek and I. Petersén, "A method for assessing alertness fluctuations from EEG spectra," *Electroencephalogr. Clin. Neurophysiol.*, vol. 55, no. 1, pp. 108–113, Jan. 1983.

[21] A. Gevins *et al.*, "Monitoring working memory load during computer-based tasks with EEG pattern recognition methods," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 40, no. 1, pp. 79–91, Mar. 1998.

[22] A. Gale, R. Davies, and A. Smallbone, "EEG correlates of signal rate, time in task and individual differences in reaction time during a five-stage sustained attention task," *Ergonomics*, vol. 20, no. 4, pp. 363–376, Jul. 1977.

[23] M. B. Sterman and C. A. Mann, "Concepts and applications of EEG analysis in aviation performance evaluation," *Biol. Psychol.*, vol. 40, nos. 1–2, pp. 115–130, May 1995.

[24] D. Popovic, M. Stikic, T. Rosenthal, D. Klyde, and T. Schnell, "Sensitive, diagnostic and multifaceted mental workload classifier (PHYSIOPRINT)," in *Proc. Int. Conf. Augmented Cognition*. Cham, Switzerland: Springer, Aug. 2015, pp. 101–111.

[25] R. H. Mogford, J. A. Guttman, S. L. Morrow, and P. Kopardekar, "The complexity construct in air traffic control: A review and synthesis of the literature," Defense Tech. Inf. Center, Fort Belvoir, VA, USA, 1995.

[26] M. L. Cummings and C. Tsonis, "Deconstructing complexity in air traffic control," *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 49, no. 1, pp. 25–29, Sep. 2005.

[27] Z. W. Pylyshyn and R. W. Storm, "Tracking multiple independent targets: Evidence for a parallel tracking mechanism," *Spatial Vis.*, vol. 3, no. 3, pp. 179–197, 1988.

[28] B. J. Scholl and Z. W. Pylyshyn, "Tracking multiple items through occlusion: Clues to visual objecthood," *Cognit. Psychol.*, vol. 38, no. 2, pp. 259–290, Mar. 1999.

[29] M. Kassner, W. Patera, and A. Bulling, "Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Computing: Adjunct Publication*, Sep. 2014, pp. 1151–1160.

[30] C. Kothe. *Lab Streaming Layer (LSL)*. [Online]. Available: https://github.com/sccn/labstreaminglayer

[31] R. J. Innes, N. J. Evans, Z. L. Howard, A. Eidels, and S. D. Brown, "A broader application of the detection response task to cognitive tasks and online environments," *Hum. Factors*, vol. 63, no. 5, pp. 896–909, 2019.

[32] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, Mar. 2004.

[33] T. T. N. Do, C. T. Lin, and K. Gramann, "Human retrosplenial theta and alpha modulation in active spatial navigation," *BioRxiv*, 2020.

[34] L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig, "ICLabel: An automated electroencephalographic independent component classifier, dataset, and website," *NeuroImage*, vol. 198, pp. 181–197, Sep. 2019.

[35] A. Delorme *et al.*, "EEGLAB, SIFT, NFT, BCILAB, and ERICA: New tools for advanced EEG processing," *Comput. Intell. Neurosci.*, vol. 2011, pp. 1–12, Oct. 2011.

[36] G. Shou, L. Ding, and D. Dasari, "Probing neural activations from continuous EEG in a real-world task: Time-frequency independent component analysis," *J. Neurosci. Methods*, vol. 209, no. 1, pp. 22–34, Jul. 2012.

[37] P. Aricò *et al.*, "Adaptive automation triggered by EEG-based mental workload index: A passive brain-computer interface application in realistic air traffic control environment," *Frontiers Hum. Neurosci.*, vol. 10, p. 539, Oct. 2016.

[38] P. Aricò *et al.*, "Human factors and neurophysiological metrics in air traffic control: A critical review," *IEEE Rev. Biomed. Eng.*, vol. 10, pp. 250–263, 2017.

[39] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," *Comput. Intell. Neurosci.*, vol. 2011, pp. 1–9, Oct. 2011.

[40] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k-means clustering algorithm," *J. Roy. Stat. Soc. C, Appl. Statist.*, vol. 28, no. 1, pp. 100–108, 1979.

[41] R. Oostenveld and T. F. Oostendorp, "Validating the boundary element method for forward and inverse EEG computations in the presence of a hole in the skull," *Hum. Brain Mapping*, vol. 17, no. 3, pp. 179–192, Nov. 2002.

[42] J. L. Lancaster *et al.*, "Automated Talairach atlas labels for functional brain mapping," *Hum. Brain Mapping*, vol. 10, no. 3, pp. 120–131, Jul. 2000.

[43] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. AE-15, no. 2, pp. 70–73, Jun. 1967.

[44] C. M. Privitera, L. W. Renninger, T. Carney, S. Klein, and M. Aguilar, "Pupil dilation during visual target detection," *J. Vis.*, vol. 10, no. 10, pp. 3–3, Aug. 2010.

[45] L. H. Negri. (2018). *PeakUtils*. Accessed: Jul. 20, 2016. [Online]. Available: https://pypi.python.org/pypi/PeakUtils

[46] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 37, no. 1, pp. 32–64, Mar. 1995.

[47] M. R. Endsley and D. J. Garland, "Pilot situation awareness training in general aviation," *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 44, no. 11, pp. 357–360, Jul. 2000.

[48] B. Lorenz, R. Parasuraman, G. R. L. Hockey, A. W. K. Gaillard, and O. Burov, "Human operator functional state in automated systems: The role of compensatory control strategies," in *Operator Functional State as a Framework for the Assessment of Performance Degradation* (NATO Science Series Sub Series I Life and Behavioural Sciences), vol. 355. 2003, pp. 224–237.

[49] R. McKendrick, T. Shaw, E. de Visser, H. Saqer, B. Kidwell, and R. Parasuraman, "Team performance in networked supervisory control of unmanned air vehicles: Effects of automation, working memory, and communication content," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 56, no. 3, pp. 463–475, May 2014.

[50] T. G. Hicks and W. W. Wierwille, "Comparison of five mental workload assessment procedures in a moving-base driving simulator," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 21, no. 2, pp. 129–143, Apr. 1979.

[51] S. Miyake, M. Hashimoto, T. Watanabe, Y. Takae, and Y. Shiraishi, "Human responses to a workload change," *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 48, no. 19, pp. 2344–2348, Sep. 2004.

[52] A. Didomenico and M. A. Nussbaum, "Interactive effects of physical and mental workload on subjective workload assessment," *Int. J. Ind. Ergonom.*, vol. 38, nos. 11–12, pp. 977–983, Nov. 2008.

[53] A. M. Basahel, M. S. Young, and M. Ajovalasit, "Impacts of physical and mental workload interaction on human attentional resources performance," in *Proc. 28th Annu. Eur. Conf. Cognit. Ergonom. (ECCE)*, 2010, pp. 215–217.

[54] M. Jiang, D. Sun, Q. Li, and D. Wang, "The usability of ventilator maintenance user interface: A comparative evaluation of user task performance, workload, and user experience," *Sci. Prog.*, vol. 103, no. 4, pp. 1–18, 2020.

[55] F. N. Biondi, A. Cacanindin, C. Douglas, and J. Cort, "Overloaded and at work: Investigating the effect of cognitive workload on assembly task performance," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 63, no. 5, pp. 813–820, Aug. 2021.

[56] P. Aricò *et al.*, "How neurophysiological measures can be used to enhance the evaluation of remote tower solutions," *Frontiers Hum. Neurosci.*, vol. 13, p. 303, Sep. 2019.

[57] G. A. Alvarez and S. L. Franconeri, "How many objects can you track?: Evidence for a resource-limited attentive tracking mechanism," *J. Vis.*, vol. 7, no. 13, p. 14, Oct. 2007.

[58] P. Cavanagh and G. A. Alvarez, "Tracking multiple targets with multifocal attention," *Trends Cognit. Sci.*, vol. 9, no. 7, pp. 349–354, 2005.

[59] J. M. Hopf, C. N. Boehler, S. J. Luck, J. K. Tsotsos, H. J. Heinze, and M. A. Schoenfeld, "Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision," *Proc. Nat. Acad. Sci. USA*, vol. 103, no. 4, pp. 1053–1058, 2006.

[60] N. G. Müller, M. Mollenhauer, A. Rösler, and A. Kleinschmidt, "The attentional field has a Mexican hat distribution," *Vis. Res.*, vol. 45, no. 9, pp. 1129–1137, Apr. 2005.

[61] Z. Pylyshyn, "The role of location indexes in spatial perception: A sketch of the FINST spatial-index model," *Cognition*, vol. 32, no. 1, pp. 65–97, Jun. 1989.

[62] Z. W. Pylyshyn, "Visual indexes, preconceptual objects, and situated vision," *Cognition*, vol. 80, nos. 1–2, pp. 127–158, Jun. 2001.

[63] Z. W. Pylyshyn, *Things and Places: How the Mind Connects With the World*. Cambridge, MA, USA: MIT Press, 2007.

[64] S. Yantis, "Multielement visual tracking: Attention and perceptual organization," *Cognit. Psychol.*, vol. 24, no. 3, pp. 295–340, Jul. 1992.

[65] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis," *Brain Res. Rev.*, vol. 29, nos. 2–3, pp. 169–195, Apr. 1999.

[66] W. Klimesch, M. Doppelmayr, H. Russegger, T. Pachinger, and J. Schwaiger, "Induced alpha band power changes in the human EEG and attention," *Neurosci. Lett.*, vol. 244, no. 2, pp. 73–76, Mar. 1998.

[67] G. F. Wilson, "An analysis of mental workload in pilots during flight using multiple psychophysiological measures," *Int. J. Aviation Psychol.*, vol. 12, no. 1, pp. 3–18, Jan. 2002.

[68] A. Capilla, J.-M. Schoffelen, G. Paterson, G. Thut, and J. Gross, "Dissociated α-band modulations in the dorsal and ventral visual pathways in visuospatial attention and perception," *Cerebral Cortex*, vol. 24, no. 2, pp. 550–561, Feb. 2014.

[69] S. Puma, N. Matton, P.-V. Paubel, É. Raufaste, and R. El-Yagoubi, "Using theta and alpha band power to assess cognitive workload in multitasking environments," *Int. J. Psychophysiol.*, vol. 123, pp. 111–120, Jan. 2018.

[70] L. R. Fournier, G. F. Wilson, and C. R. Swain, "Electrophysiological, behavioral, and subjective indexes of workload when performing multiple tasks: Manipulations of task difficulty and training," *Int. J. Psychophysiol.*, vol. 31, no. 2, pp. 129–145, Jan. 1999.

[71] M. E. Smith, A. Gevins, H. Brown, A. Karnik, and R. Du, "Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 43, no. 3, pp. 366–380, Sep. 2001.

[72] S. H. Fairclough, L. Venables, and A. Tattersall, "The influence of task demand and learning on the psychophysiological response," *Int. J. Psychophysiol.*, vol. 56, no. 2, pp. 171–184, May 2005.
[73] S. H. Fairclough and L. Venables, "Prediction of subjective states from psychophysiology: A multivariate approach," *Biol. Psychol.*, vol. 71, no. 1, pp. 100–110, Jan. 2006.
[74] K. Ryu and R. Myung, "Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic," *Int. J. Ind. Ergonom.*, vol. 35, no. 11, pp. 991–1009, Nov. 2005.
[75] T. Ota, R. Toyoshima, and T. Yamauchi, "Measurements by biphasic changes of the alpha band amplitude as indicators of arousal level," *Int. J. Psychophysiol.*, vol. 24, nos. 1–2, pp. 25–37, Nov. 1996.
[76] T. Harmony *et al.*, "EEG delta activity: An indicator of attention to internal processing during performance of mental tasks," *Int. J. Psychophysiol.*, vol. 24, nos. 1–2, pp. 161–171, 1996.
[77] O. Jensen and C. D. Tesche, "Frontal theta activity in humans increases with memory load in a working memory task," *Eur. J. Neurosci.*, vol. 15, no. 8, pp. 1395–1399, Apr. 2002.
[78] J. Jacobs, G. Hwang, T. Curran, and M. J. Kahana, "EEG oscillations and recognition memory: Theta correlates of memory retrieval and decision making," *NeuroImage*, vol. 32, no. 2, pp. 978–987, Aug. 2006.
[79] W. Klimesch, "Memory processes, brain oscillations and EEG synchronization," *Int. J. Psychophysiol.*, vol. 24, nos. 1–2, pp. 61–100, 1996.
[80] P. Sauseng, B. Griesmayr, R. Freunberger, and W. Klimesch, "Control mechanisms in working memory: A possible function of EEG theta oscillations," *Neurosci. Biobehav. Rev.*, vol. 34, no. 7, pp. 1015–1022, Jun. 2010.
[81] S. M. Slobounov, K. Fukada, R. Simon, M. Rearick, and W. Ray, "Neurophysiological and behavioral indices of time pressure effects on visuomotor task performance," *Cognit. Brain Res.*, vol. 9, no. 3, pp. 287–298, Jun. 2000.
[82] I. Batmaz and M. Ozturk, "Using pupil diameter changes for measuring mental workload under mental processing," *J. Appl. Sci.*, vol. 8, no. 1, pp. 68–76, Dec. 2007.
[83] T. Kosch, M. Hassib, D. Buschek, and A. Schmidt, "Look into my eyes: Using pupil dilation to estimate mental workload for task complexity adaptation," in *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, Apr. 2018, pp. 1–6.
[84] M. Á. Recarte, E. Pérez, Á. Conchillo, and L. M. Nunes, "Mental workload and visual impairment: Differences between pupil, blink, and subjective rating," *Spanish J. Psychol.*, vol. 11, no. 2, pp. 374–385, 2008.
[85] A. Widyanti, N. F. Sofiani, H. R. Soetisna, and K. Muslim, "Eye blink rate as a measure of mental workload in a driving task: Convergent or divergent with other measures?" *Int. J. Technol.*, vol. 8, no. 2, p. 283, Apr. 2017.
[86] B. Mehler, B. Reimer, and Y. Wang, "A comparison of heart rate and heart rate variability indices in distinguishing single-task driving and driving under secondary cognitive workload," in *Proc. 6th Int. Driving Symp. Hum. Factors Driver Assessment, Training Vehicle Design*, 2011, pp. 1–8.
[87] T. Heine, G. Lenis, P. Reichensperger, T. Beran, O. Doessel, and B. Deml, "Electrocardiographic features for the measurement of drivers' mental workload," *Appl. Ergonom.*, vol. 61, pp. 31–43, May 2017.
[88] A. F. Kramer, "Physiological metrics of mental workload: A review of recent progress," in *Multiple-Task Performance*, D. L. Damos, Ed. London, U.K.: Taylor & Francis, 1991, pp. 279–328.
[89] D. A. Valentino, J. E. Arruda, and S. M. Gold, "Comparison of QEEG and response accuracy in good vs poorer performers during a vigilance task," *Int. J. Psychophysiol.*, vol. 15, no. 2, pp. 123–133, Sep. 1993.
[90] S. Athènes, P. Averty, S. Puechmorel, D. Delahaye, and C. Collet, "ATC complexity and controller workload: Trying to bridge the gap," in *Proc. Int. Conf. HCI Aeronaut.*, Cambridge, MA, USA, Oct. 2002, pp. 56–60.
[91] J. C. Sperandio, "Variation of operator's strategies and regulating effects on workload," *Ergonomics*, vol. 14, no. 5, pp. 571–577, Sep. 1971.