Neuroergonomics and Cognitive Engineering, Vol. 102, 2023, 171–181
https://doi.org/10.54941/ahfe1003017

Mental Workload Classification During Simulated Flight Operations Based on Cardiac and Neural Dynamics Recorded Using the MUSE 2 Low-Cost System

Frédéric Dehais[1,2], Simon Ladouce[1], Juan Jesus Torre Tresols[1], Ludovic Darmet[1], and Daniel Callan[1,3]

[1]ISAE-SUPAERO, Department of Aerospace Vehicles Design, Université Fédérale Midi-Pyrénées, Toulouse, France
[2]Drexel Biomedical Engineering, Drexel University, Philadelphia, PA, USA
[3]Neural Information Analysis Laboratories, Advanced Telecommunications Research Institute International, Kyoto, Japan

ABSTRACT

The advancement of low-cost and highly portable physiological systems presents promising opportunities for monitoring human cognitive processes during daily-life activities and more complex tasks such as operating an aircraft. The Muse 2 system combines electroencephalography (EEG) and photoplethysmography (PPG) sensors allowing the extraction of neural dynamics features in the time and frequency domains and heart rate. In a study, we equipped five pilots with the Muse 2 system while they performed a low-load and high-load traffic pattern task along with a passive auditory oddball task. The group-level analyses revealed that participants exhibited higher average heart rate, lower power spectrum density in the alpha band, decreased P300 amplitude in the high-load compared to the low-load condition. These results are in line with previous laboratory research conducted in highly controlled settings and research-grade instrumentations. The classification of the two levels of mental workload reached 93.2% accuracy on a single-trial basis based on EEG frequency features. Post-hoc analysis revealed that the classifier mainly relied on motion artefact features in the beta and gamma bands. The classifiers using heart rate and ERPs features reached 76% and 77.8% classification accuracy, respectively. Despite its interest, this system presents some limitations for mobile and neuroergonomics applications notably with regards to the limited number of electrodes preventing the use of advanced signal processing techniques to address noise and artifacts in the signals.

Keywords: Mental workload, Low-cost EEG system, Muse 2, Classification, Neuroergonomics

INTRODUCTION

Pilots are subjected to high levels of cognitive effort during flight, which can adversely affect their performance and safety (Callan et al., 2018). Currently, the assessment of mental workload is typically based on subjective measures such as self-report questionnaires. These methods have limitations, including low reliability and the inability to provide real-time data. Behavioral measures provide direct measures of performance that can be leveraged for the implementation of assistive Human-Machine Interfaces approaches. These measures may however not be recorded during flight segments such as the cruising phase, where very few actions are performed. Using objective measures of mental workload derived from physiological signals would allow for circumventing these limitations (Fairclough et al., 2005). For instance, heart rate and its variability (Scannella et al., 2018) have been used as proxies to assess mental efforts.

There has also been considerable interest in Neuroergonomics research to use brain measures like electroencephalography (EEG) as a more direct way to probe the deployment of human operator's cognitive resources (Fairclough et al, 2020, Dehais et al, 2020a). A classical approach has been to extract specific frequency-domain features to account for mental effort (Borghini et al., 2014, Belkhiria et al, 2020). Alternatively, time-domain analyses of EEG signals can also be used to detect transient changes in mental workload using probing techniques. For instance, the volunteers' electrophysiological response to an auxiliary auditory oddball task, in which are presented among frequent standard distractors, is assessed. The rare/target auditory stimuli are known to evoke a positive deflection around 300 *ms* (the "P300" Event-Related Potential). The amplitude of the P300 has been shown to index the resources allocated to the selective attentional processing of task-relevant information. Following a principle of limited pool of attentional resources, the amplitude of the P300 is adversely affected when cognitive resources have to be distributed across different tasks performed jointly. As such, P300 amplitude can be used as an indirect index of the level of mental workload as its amplitude is assumed to decrease when the primary task demands increase (Roy et al. 2016a, Roy et al. 2016b, Brouwer et al., 2012). Most of the previous studies were conducted with costly and bulky wet-electrodes research-grade EEG systems that prevent their use for everyday operation (Somon et al., 2022b). As sensors become ever more portable and less intrusive, they offer new perspectives for neuroergonomics applications (Dehaisb et al., 2020). For instance, several studies in simulated and real-flight conditions have successfully shown the potential of wireless dry EEG systems to detect critical operators' states (Dehais et al., 2020b, Dehais et al., 2018, Callan et al., 2018, Scholl et al., 2016) even with very few electrodes (Dehais et al, 2022, Somon et al., 2022a, Dehais et al., 2019, Getzmann, 2021).

In recent years, the development of neurotechnological hardware has led to an expanding range of low-cost devices, such as the Emotive, Dreem, OpenBCI and the Muse system, aiming at everyday life applications (health, entertainment, domotic...) The Muse-2 system comprises five EEG electrodes (including a reference electrode) and a photoplethysmography (PPG) sensor enabling an optical-based measure of heart rate. Several studies indicated the potential of these unobtrusive devices for mobile research (Cannard, Wahbeh & Delorme, 2021, Krigolson et al., 2021, Krigoloson et al., 2017) and the development of brain-computer interface (BCI) as shown by Simar et al., (2020). These studies were however conducted in controlled laboratory environments and during tasks where subjects' movements were limited. Therefore, there is a need to benchmark these sensors in applied contexts where users are freely engaged in ecologically valid tasks without any restriction on body movements, such as in a flight simulator. To do so, we replicated the study from Dehais et al. (2019) using a 6-dry electrode system in which pilots had to perform a traffic pattern routine under two load conditions alongside an auditory oddball task. Our goal is to demonstrate that ERPs, frequency features, as well as heart rate derived from the Muse-2 system, can be used to perform mental workload single trial classification for BCI purposes.

MATERIAL AND METHODS

Participants

Five male participants (mean age = 25 years), all students from the Aeronautical & Space Superior Institute (ISAE-SUPAERO) - Toulouse Federal University, took part in this experiment. The experiment was approved by the Ethical Committee (CER) of Toulouse Federal University (approval number 2022-527).

Scenario

We used the ISAE-SUPAERO flight simulator to conduct the experiments. The simulator comprises 8 external displays that present a panoramic rendering of the environment rendered through the Flight-Gear open-source software. The scenario consisted of two consecutive traffic patterns that were performed at Blagnac airport virtual location. In one traffic pattern, defined as the low load condition, the participants (left-seated) were monitoring the flight operated by the flight instructor (right-seated). In the other traffic pattern, defined as the high load condition, the participants were operating the aircraft themselves while being supervised by the flight instructor. Each traffic pattern lasted for approximately 10 minutes and 40 seconds. Along with the flying tasks (i.e. monitoring and flying), the participants were instructed to perform a passive auditory oddball paradigm comprising a total of 250 auditory stimuli: 25 % were targets (50 normalized pure tones at 1100 Hz) and 75% were non-targets (200 normalized pure tones at 1000 Hz). The inter-trial interval was set to 2000 ms with a 1000-ms jitter window. EEG data were streamed using Lab Streaming Layer (LSL) and the Muse-LSL library (Barachant et al., 2019). The oddball task was implemented in Matlab and markers corresponding to the onsets of stimulus are also streamed on LSL to enable synchronization with EEG data. After each landing, participants were asked to report the number of rare auditory sounds. The order of the scenarios was counterbalanced across subjects to control for fatigue and training effects.

Subjective and Behavioural Measures

Participants were asked to report their level of workload on an analog scale (0 = low, 10 = high). Task performance was measured as the ratio of auditory targets reported by the participant after completing each traffic pattern and the actual number of target stimuli played over the corresponding timeframe.

EEG: Event Related Potential

EEG data were preprocessed using the Matlab EEGLab toolbox v.2021. The continuous EEG data were filtered between 0.1-20 Hz (windowed-sinc FIR filter with an order of 250). The continuous EEG signals were epoched around target markers (timestamps were provided by the stimulus presentation software through LSL) from 0.2 s prior to 1 s after stimulus onset. The outlier epochs were rejected using the *pop_jointprob* EEGlab function (standard deviation threshold set to 2 for the *locthresh* and *globthresh* parameters). Lastly, the epochs were baseline normalized using data from 200 to 0 *ms* prior to stimulus onset. Further analyses and classification were performed using the Python-based library scikit learn. The classification used a traditional pipeline for single-trial ERP classification (Barachant et Congedo, 2014). This pipeline was implemented using Pyriemann for the covariance feature extraction and Scikit-Learn for the classifiers. First, the dimensionality was reduced using a supervised spatial filter: Xdawn with 4 filters. Xdawn enhances the target response with respect to the non-target response for ERP (B. Rivet et al., 2009).

Then for each class, a prototyped response **P** was obtained by average across trials. For each trial $\mathbf{X}_i$, a super trial $\mathbf{S}_i$ was built using the concatenation of **P** and the trial $\mathbf{X}_i$. These super trials $\mathbf{S}_i$ were then used for covariance estimation. The trials were converted into covariance matrices to take into account the spatial structure of the signal (Barrachant & Congedo, 2014). Next, the covariance matrices were projected into their tangent space, using the geometric mean of all the covariance matrices as a reference point (Barachant et al., 2012). After this projection, each covariance matrix was represented by a vector upon which a logistic regression (no regularization) was applied for classification. The performance was evaluated using balanced accuracy and a 10-fold cross-validation. The folds were stratified to ensure that each of them contained the same proportion of target and non-target trials. We reported balanced accuracy measures since the number of epochs between the two mental workload conditions could slightly differ.

EEG: Frequency Features

Features extraction was conducted using the Python toolbox MNE (Gramfort et al., 2013), while the classification and cross-validation parts are processed using Scikit-Learn. Firstly, we have computed the average power spectral density (PSD) in the alpha band ([8 12] Hz) on frontal electrodes at the group level for descriptive statistics purposes. All the averaged PSDs in this study were computed using the Welch method on sliding windows of length 256. Sub-bands PSD data were converted in dB to take the 1/f law of EEG signal spectral power into account and to reduce the contribution of large values, such as outliers (Lotte, 2014).

Thereafter for the individual single trial classification pipeline, sub-bands PSD metrics were extracted from non-overlapping 3s windows. The approach aimed to predict whether an unseen 3s segment of EEG data the pilot was monitoring or flying. For each window, 5 spectral features were extracted. These features corresponded to the averaged PSD (in dB) of 5 frequency bands of interest: delta [1 4] Hz, theta [4 8] Hz, alpha [8 12] Hz, beta [12 30] Hz and gamma [30 45] Hz). Hence, for each subject, we obtained a matrix with 5 columns corresponding to the 5 spectral features and rows corresponding to the different 3s windows (440 on average), both from the flying and monitoring conditions. The corresponding windows labels were 1 if the subject was flying and 0 if monitoring.

A specific classifier was trained for each subject, using only his EEG data. As in Dehais et al. (2019), the classification approach used Linear Discriminant Analysis (LDA) preceded by a standardization of features (mean centering and scaling to unit variance). We implemented a nested cross-validation: an outer loop with 5-fold cross-validation to split between train and test for performance evaluation and another inner-loop of again 5-fold cross-validation for train and validation split. Validation data were used to find optimal hyper-parameters, individually for each subject, for the LDA (solver: 'lsqr' or 'eigen' and number of components: 1, 5, 10). The nested cross-validation was repeated 5 times, with different seeds leading to a total of 25 performance measures for each subject that were then averaged and reported. We also inspected the absolute weightvalues of the LDA models (on the first component) reflecting the importance attributed to features as a proxy to measure their contribution to the model. We computed four different pipelines: using all spectral features, using only beta and gamma bands, using only delta and theta bands, and using only alpha band activity.

PPG

The Muse-2 has three PPG channels (ambient, green and red lights). As recommended by Maeda et al. (2008), the green light PPG channel (#2) was preferred over the infrared one since it correlates better with ECG signal. We used the neurokit2[1] Python package (Makowski et al., 2021) dedicated to compute the HR derived from the PPG sensor. The data for the high (PF) and low (PM) mental workload conditions were cleaned with the *ppg_process* function (default settings) - see Figure 1. and the PPG HR was then extracted using the *ppg_analyze* function (default settings). From the HR, we have extracted the mean and standard deviation on 3-second non-overlapping windows. These two features are then supplied to an LDA classifier to discriminate between the two experimental conditions: pilot flying and monitoring.

RESULTS

Descriptive Statistics

The following Table 1 reports the descriptive findings at the group level for the subjective, behavioral and physiological measures.

Statistical analyses of the ERPs were carried out using a 2-way repeated measure analysis of variance (ANOVA) with load (Low, High) and Electrodes

[1]https://neuropsychology.github.io/NeuroKit. © Copyright 2020

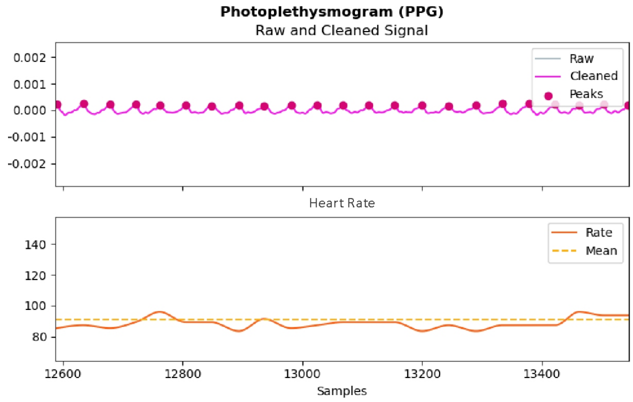

Figure 1: Example of PPG data processed with Neurokit2 for peaks detection (above) and HR estimation (below).

Table 1. Subjective, behavioural and neurophysiological findings (Folds et al. 2008).

Metrics	Low load (PM) Mean (SD)	High load (PF) Mean (SD)
Subjective Mental workload	4.0 (1.6)	7.2 (0.8)
Absolute number of erroneous auditory targets	4.4 (4.3)	10 (5.7)
HR PPG (bpm)	88.8 (20.1)	93.1 (17.2)
Frontal Alpha (dB)	109.9 (2.8)	111.3 (2.3)

(TP9 and TP10) as within-subject factors. We found a load condition × electrodes significant interaction ($p < 0.01$, FDR corrected). This effect was due to higher P300 amplitude for the target sound in the low load compared to the high load condition on TP9 and TP10 (see Figure 2).

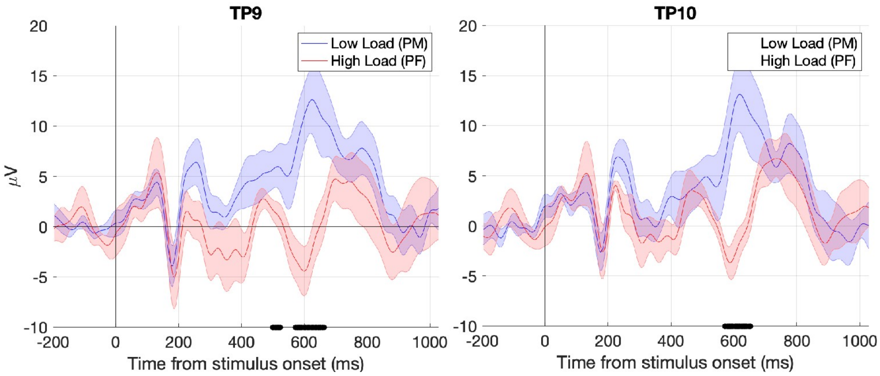

Figure 2: Grand averaged waveforms of the ERPs for temporal electrodes with standard error (shapes). Blue: Low load conditions for auditory targets. Red: High load condition for auditory targets. The black lines on the x axis specify the time range when the target sound-related and the frequent sound-related ERP amplitudes were significantly different ($p < 0.01$).

Classification Accuracy for the Different Pipelines

Table 2
presents the classification accuracy for the different combinations of EEG frequency features to classify the low (PM) vs high workload (PF) scenario.

Table 2. Classification accuracy for the EEG frequency pipeline.

All features Mean SD	Delta & Theta Mean (SD)	Alpha Mean (SD)	Beta Mean (SD)& Gamma
% (2.5)	61.0 % (5.1)	50.9% (5.2)	91.6% (2.5)

The mean classification accuracy reached 75.8% (SD: 18.1%) for the PPG based classifier and 77.8% (SD: 7.6%) for the ERP based classifier (see Figure 3).

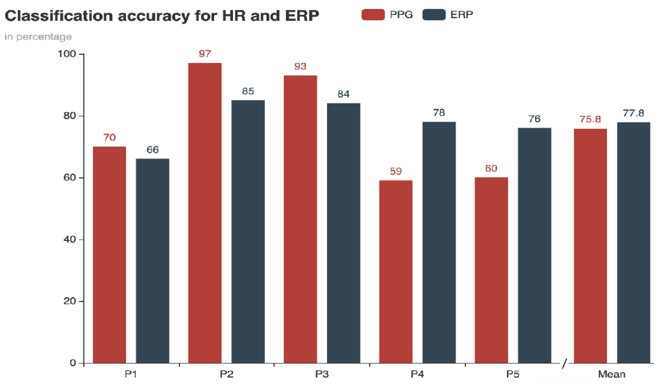

Figure 3: Mean classification accuracy for the five participants with the PPG (in red) and ERP (in blue) based classifiers.

DISCUSSION

The objective of this study was to show the feasibility to measure pilot's mental workload with a low cost highly mobile EEG/PPG system. The subjective and behavioral measures confirmed that we successfully manipulated mental demand in the two scenarios. The participants reported higher mental workload and committed more counting errors on the auditory task when flying than when they were only monitoring the flight. The participants also exhibited a higher mean PPG HR, a lower mean alpha band power during the performance of the flying condition compared to the monitoring condition. These trends are consistent with the literature since higher mental demand is known to increase cardiac activity (higher HR) and arousal (lower PSD in the alpha band). However, these preliminary findings have to be taken carefully since our sample size was very small (N = 5) and the standard deviations for these metrics were high. Interestingly enough, the time-domain analysis showed that the P300 amplitude for the auditory target was statistically higher in the low load condition (PM) than in the high load condition on the two temporal electrodes. These results are in line with previous findings (Dehais et al., 2019, Roy et al. 2016a, Roy et al. 2016b, Brouwer et al., 2012) who demonstrated that the amplitude of the evoked P300 (used as a probe of available attentional resources) was negatively affected when primary task difficulty increased, leaving less cognitive resources to process the secondary auditory tasks. Recent works have also demonstrated the processing of visual and inertial flow of information recapture attentional resources and therefore have adverse effects on attentional resources allocated to a secondary auditory task (Ladouce et al., 2019). However, it can't be excluded that the absence of P300 in the high load might also be due to muscular artefacts (neck, eye movement, arms) as a consequence of operating the thrust and joystick.

Our main objective was to perform single trial classification over the different signals (EEG and PPG) to discriminate the two load conditions. The classification accuracy was very high for the frequency feature-based classifier (92.5%) when combining all the features. However subsequent analyses revealed the strong influence of the beta and gamma band-based features on this classification score (see Table 2). Whitham et al. (2007) demonstrated that EEG frequencies above 20 Hz are highly contaminated by electromyographic activity. In our experiment, the pilots exhibited higher motor activity in the flying condition thus biasing the classifier that gave more importance to these high frequency features. The theta-delta bands-based classifier reached 61% of accuracy. It is above chance level, however, one has to keep in mind that these frequencies are also contaminated by blinks (Kaia et al., 2020). When using only the alpha band, a frequency that reflects endogenous brain rhythm, the classification was at chance level. One reason for this low accuracy is that it is generally recommended to compute alpha band PSD on parietal sites (Arico et al., 2015, Erwin et al., 2016, Fairclough et al., 2005) rather than on temporal and prefrontal ones as we did because of the Muse-2 electrodes location. Besides the muscular artifacts in the EEG signal, the workload was not constant during the 10 minutes of each scenario (i.e., always low in PM condition vs always high in PF condition). Scannella et al. (2018) reported that the landing and take-off legs led to higher mental demand than the downwind one. As a consequence, a portion of the labels used to train the classifier (i.e. label 1 for all epochs of PF and label 0 for all epochs of PM) may not reflect the actual experienced workload by the pilots. One could imagine that the experienced workload for some epochs extracted during the downwind phase of PF was actually closer to the low workload of the PM scenario, than the very intense workload of landing and take-off legs during PF. This portion of fuzzy labels is hard to estimate but probably significant and can partially explain the poor testing accuracy obtained by the pipeline using only the alpha band power feature. In other terms, this binary segmentation therefore falls short in reflecting how the mental workload induced by a cognitive task is not monotonic/stationary but rather fluctuates over transient episodes throughout a task. Interestingly enough the classification pipeline using data locked around ERPs of the auxiliary oddball task has provided very good performance (77.8%) with relatively low variance. The classification pipeline was using only data from four electrodes (and did not target central and parietal sites). It is far better than the ERPs based classifier implemented by Dehais et al (2019) with a 6-dry EEG system that did not exceed chance level. One of the reasons is that here, we used advanced machine learning techniques combining the spatial filtering of Xdawn with Riemannian based classification algorithms that have demonstrated high classification scores even in ecological settings (Somon et al., 2022a). As outlined above, the experienced workload fluctuated more than the binary labels used by the classification. However, the reduction on average of oddball ERPs amplitude during the PF seems to indicate that the transient variations of workload have less impact on the modulation of ERPs level. This could explain why the ERPs based pipeline performed better than the spectral pipeline that excludes bands with muscular artifacts. This is in line with the claim from Roy et al. (2016) that ERPs based classifiers produce a more stable classification over time compared to frequency base classification pipeline. However, a drawback of the ERP-based pipeline is that it requires the use of additional non-natural stimuli, the oddball, that might not be suited for real flight operations. Eventually, the HR-PPG based classifier also performed reasonably well on average (75.8%) but this performance was characterized by a much higher variance than the ERP-based classification.

CONCLUSION

This experiment, together with previous studies (Cannard, Wahbeh & Delorme, 2021, Krigolson et al., 2021, Krigoloson et al., 2017) confirms that a low-cost and portable device offers interesting prospects for research. The quick and easy setup of both PPG and EEG sensors are appealing features within the context of applied research. The reduced number of dry electrodes however comes at the expense of signal-to-noise ratio and preventing from performing state of the art EEG cleaning procedure that require high number of channels (e.g., ICA). As a consequence, our results suggest that ERPs based classifiers appeared to offer the best trade-off in terms of accuracy and responsiveness for estimating mental workload.

ACKNOWLEDGMENT

The authors would like to acknowledge AXA (Neuroerogonomics chair) research fund, ANITI (Neurotechnology chair) and AID (Power brain project) for supporting this study and all the participants who took part in this study.

REFERENCES

Arico, P., Borghini, G., Di Flumeri, G., Colosimo, A., Graziani, I., Imbert, J. P.,... & Babiloni, F. (2015, August). Reliability over time of EEG-based mental workload evaluation during Air Traffic Management (ATM) tasks. In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 7242–7245). IEEE.

Barachant, A., & Congedo, M. (2014). A plug & play P300 BCI using information geometry. *arXiv preprint arXiv:1409.0107*.

Barachant, A. Morrison, D., Banville, H., Kowaleski, J, Shaked, U., Chevallier, S., & Torre Tresols, J. J.. (2019, May 25). muse-lsl (Version v2.0.2). Zenodo. http://doi.org/10.5281/zenodo.3228861

Belkhiria, C., & Peysakhovich, V. (2020). Electro-encephalography and electro-oculography in aeronautics: A review over the last decade (2010–2020). Frontiers in Neuroergonomics, 1, 606719.

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., & Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and cardrivers for the assessment of mental workload, fatigue and drowsiness. Neuroscience & Biobehavioral Reviews, 44, 58–75.

Brouwer, A. M., Hogervorst, M. A., Van Erp, J. B., Heffelaar, T., Zimmerman, P. H., & Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. Journal of neural engineering, 9(4), 045008.

Callan, D. E., Gateau, T., Durantin, G., Gonthier, N., & Dehais, F. (2018). Disruption in neural phase synchrony is related to identification of in attentional deafness in real world settings. Human brain mapping, 39(6), 2596–2608.

Cannard, C., Wahbeh, H., & Delorme, A. (2021). Validating the wearable-MUSE headset for EEG spectral analysis and Frontal Alpha Asymmetry. In 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 3603–3610).

Dehais, F., Duprès, A., Blum, S., Drougard, N., Scannella, S., Roy, R. N., & Lotte, F. (2019). Monitoring pilot's mental workload using ERPs and spectral power with a six-dry-electrode EEG system in real flight conditions. Sensors, 19(6), 1324.

Dehais, F.; Duprès, A.; Di Flumeri, G.; Verdière, K. J.; Borghini, G.; Babiloni, F.; Roy, R. N. (2018) Monitoring pilot's cognitive fatigue with engagement features in simulated and actual flight conditions using a hybrid fNIRS-EEG passive BCI. IEEE SMC 2018, 544–549.

Dehais, F., Karwowski, W., & Ayaz, H. (2020a). Brain at work and in everyday life as the next frontier: grand field challenges for neuroergonomics. Frontiers in Neuroergonomics, 1, 583733.

Dehais, F., Lafont, A., Roy, R., & Fairclough, S. (2020b). A neuroergonomics approach to mental workload, engagement and human performance. Frontiers in neuroscience, 268.

Dehais, F., Somon, B., Mullen, T., & Callan, D. E. (2020c). A neuroergonomics approach to measure pilot's cognitive in capacitation in the real world with EEG. In Advances in Neuroergonomics and Cognitive Engineering: Proceedings of the AHFE 2020 conference, July 16-20, 2020, USA (pp. 111–117). Springer International Publishing.

Dehais, F., Ladouce, S., Darmet, L., Tran-Vu, N., Ferraro, G., Tresols, J. T.,... & Labedan, P. (2022). Dual passive reactive brain computer interface: a novel approach to human-machine symbiosis. *Frontiers in Neuroergonomics*.

Fairclough, S. H., Venables, L., & Tattersall, A. (2005). The influence of task demand and learning on the psychophysiological response. International Journal of Psychophysiology, 56(2), 171–184.

Fairclough, S. H., & Lotte, F. (2020). Grand challenges in neurotechnology and system neuroergonomics. Frontiers in Neuroergonomics, 2.

Getzmann, S., Reiser, J. E., Karthaus, M., Rudinger, G., & Wascher, E. (2021). Measuring correlates of mental workload during simulated driving using cee-grid electrodes: a test–retest reliability analysis. Frontiers in neuroergonomics, 2, 729197.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C.,... & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. Frontiers in neuroscience, 267.

Kaya I. (2020) Brief Summary of EEG Artifact Handling. ArXiv10.13140/RG.2.2.25108.04484

Krigolson, O. E., Williams, C. C., Norton, A., Hassall, C. D., &Colino, F. L. (2017). Choosing MUSE: Validation of a low-cost, portable EEG system for ERP research. Frontiers in neuroscience, 11, 109.

Krigolson, O. E., Hammerstrom, M. R., Abimbola, W., Trska, R., Wright, B. W., Hecker, K. G., & Binsted, G. (2021). Using Muse: Rapid mobile assessment of brain performance. Frontiers in Neuroscience, 15, 634147.

Ladouce, S., Donaldson, D. I., Dudchenko, P. A. et al. Mobile EEG identifies the re-allocation of attention during real-world activity. Sci Rep 9, 15851 (2019). https://doi.org/10.1038/s41598-019-51996-y

Lotte, F. A tutorial on EEG signal-processing techniques for mental-state recognition in brain-computer interfaces (2014). In Guide to Brain-Computer Music Interfacing; Springer: Berlin, Germany ; pp. 133–161

Maeda, Y., Sekine, M., Tamura, T., Moriya, A., Suzuki, T., & Kameyama, K. (2008, August). Comparison of reflected green light and infrared photoplethysmography. In 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 2270–2272). IEEE.

Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H.,... & Chen, S. H. (2021). NeuroKit2: A Python toolbox for neurophysiological signal processing. Behavior research methods, 53(4), 1689-1696.

Rivet, B., Souloumiac, A., Attina, V., Gibert., G. xDAWN algorithm to enhance evoked potentials: application to brain-computer interface. *IEEE Transactions on Biomedical Engineering*, 2009, 56 (8), pp. 2035–43.

Roy, R. N., Bonnet, S., Charbonnier, S., & Campagne, A. (2016a). Efficient workload classification based on ignored auditory probes: a proof of concept. Frontiers in human neuroscience, 10, 519.

Roy, R. N., Charbonnier, S., Campagne, A., & Bonnet, S. (2016b). Efficient mental workload estimation using task-independent EEG features. Journal of neural engineering, 13(2), 026019.

Scannella, S., Peysakhovich, V., Ehrig, F., Lepron, E., & Dehais, F. (2018). Assessment of ocular and physiological metrics to discriminate flight phases in real light aircraft. Human factors, 60(7), 922–935.

Scholl, C. A.; Chi, Y. M.; Elconin, M.; Gray, W. R.; Chevillet, M. A.; Pohlmeyer, E. A. (2016) Classification of pilot-induced oscillations during in-flight piloting exercises using dry EEG sensor recordings. In Proceedings of the IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 17–20 August 2016; pp. 4467–4470.

Simar, C., Petieau, M., Cebolla, A., Leroy, A., Bontempi, G., & Cheron, G. (2020, July). EEG-based brain-computer interface for alpha speed control of a small robot using the MUSE headband. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1–4). IEEE.

Somon, B., Giebeler, Y., Darmet, L., & Dehais, F. (2022a). Benchmarking cEEGrid and solid gel-based electrodes to classify in attentional deafness in a flight simulator. Frontiers in Neuroergonomics, 2.

Somon, B., Roy, R. N., Simonetti, I., & Dehais, F. (2022b). Ecological measures of cognitive impairments in aeronautics: theory and application. In Current Research in Neuroadaptive Technology (pp.117-138). Academic Press.

Whitham, E. M., Pope, K. J., Fitzgibbon, S. P., Lewis, T., Clark, C. R., Loveless, S.,... & Willoughby, J. O. (2007). Scalp electrical recording during paralysis: quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. Clinical neurophysiology, 118(8), 1877-1888.