Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 1025–1036
https://doi.org/10.54941/ahfe1008333

Differential Sensitivity of EEG, ECG, Respiration, and GSR to Mental States Classification: A Comparative Machine-Learning Analysis

Francis Xian Logah[1], Mawuli Agboklu[1], Younho Seong[1], Sun Yi[1], and Joseph Fagerburg[2]

[1]North Carolina A&T State University, Greensboro, NC, 27411, United States
[2]BTI Software Engineering, United States

ABSTRACT

Understanding a pilot's cognitive state is critical for improving flight safety, optimizing human performance, and enabling intelligent adaptive cockpit systems. Although electroencephalography (EEG) has long been considered the benchmark for measuring cognitive workload, there has been a growing interest in the ability of other physiological signals, such as electrocardiography (ECG), respiration, and galvanic skin response (GSR) to reliably detect changes in mental state either individually or combined. This study uses four supervised machine learning algorithms: Logistic Regression, Support Vector Machine (SVM), Random Forest, and Multilayer Perceptron (MLP) to compare the sensitivity of these physiological modalities for classifying pilot mental states and examines whether multimodal physiological fusion provides superior performance. Model performance was assessed using Macro-F1 score, confusion matrices, and state-specific recall. The study showed that fusing multiple physiological signals consistently produced the most accurate mental-state classification. Multimodal fusion achieved the highest overall performance with Logistic Regression (Macro-F1 = 0.973); outperforming every individual single modality. Among the individual physiological modalities, MLP achieved the highest EEG performance (Macro-F1 = 0.930). More importantly, the fusion of ECG, Respiration, and GSR substantially improved classification accuracy with a maximum Macro-F1 score of 0.902 using the SVM classifier. This represents an approximately 18–37% point increment over the individual peripheral modalities. Overall, the study highlights the value of multimodal physiological sensing for mental state assessment. This work supports the development of robust, real-time pilot monitoring systems capable of enhancing adaptive cockpit technologies, decision-support systems, and human-centered aviation operations.

Keywords: Electroencephalography, Multimodal, Multilayer perceptron, Algorithm, Galvanic, Physiological, Machine-learning, Fusion, Respiration

INTRODUCTION

Technology in the field of aviation is shifting focus on workload from being physical to being more cognitive. Relying on more automated systems means pilots have to do more aircraft system monitoring, interpret multiple sources

Received September 24, 2026; Revised October 23, 2026; Accepted November 1, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

of information, maintain situational awareness, manage unexpected events, and intervene when automated systems are unable to adequately address changing operational conditions. This puts more strain on the cognitive abilities of pilots, making it necessary to evaluate the mental states of pilots. That notwithstanding, the appropriate modality and machine-learning algorithm to use also becomes a challenge.

Traditional approaches in mental state evaluations have relied more on subjective ratings such as the NASA-TLX. However, their inability to continuously track changes in the mental state of an operator in real-time makes them undesirable. To solve this problem, physiological measures have been employed to provide continuous and relatively objective indications of changes occurring during task performance. Among the physiological measures, the Electroencephalography (EEG) has been used extensively in cognitive studies. However, research has shown that EEG signals are susceptible to noise (Alreshidi, Moulitsas, and Jenkins, 2023). As a contingency, peripheral physiological measures such as the Galvanic Skin Response (GSR), Electrocardiogram (ECG), Respiration, etc. have also been identified as candidates for mental state studies. However, the physiological differences among the modalities bring into question their relative sensitivity to different mental states. The introduction of different machine-learning techniques in this study offers the opportunity to leverage on the unique strengths of each algorithm to perform a direct comparison between the performance of each algorithm on the different physiological modalities.

The aim of this study is to compare the sensitivity of Electroencephalography (EEG), Galvanic Skin Response (GSR), Electrocardiogram (ECG) and the Respiration Rate in evaluating the mental state of a pilot during the performance of flight simulations and under different stress conditions. Four machine learning algorithms: Logistic Regression, Support Vector Machine (SVM), Random Forest, and Multilayer Perceptron (MLP) are evaluated across six physiological configurations: EEG, ECG, respiration, GSR, peripheral fusion (ECG + respiration + GSR), and complete multimodal fusion (EEG + ECG + respiration + GSR). Classification performance was then evaluated using Macro-F1 score and recall.

This study contributes to the field of neuroergonomics research by determining not only which physiological modality provides the strongest overall classification, but also which modalities are most sensitive to particular mental states and whether their integration provides complementary benefits. This has potential implications for sensor selection and the development of adaptive cockpit technologies, and real-time pilot state monitoring.

RELATED WORKS

Cognitive Workload and Physiological Assessment

Human performance critically depends on cognitive workload; the mental effort required to process information, make decisions, and execute actions under varying operational demands (Tierney et al., 2025). Excessive cognitive workload can degrade performance, elevate error rates, and compromise safety. Traditionally, workload has been inferred through subjective ratings

such as the NASA-TLX (Dias et al., 2026). However, these approaches lack temporal precision and often fail to capture rapid fluctuations in mental state during dynamic operations. In response to these limitations, physiological measures of mental state including electroencephalography (EEG), electrocardiography (ECG), respiration patterns, and galvanic skin response (GSR) (Alshehhi et al., 2026) have emerged as objective indicators of autonomic nervous system engagement and cognitive demand. Physiological measurement is grounded in psychophysiological stress theories which posits that cognitive demand modulates autonomic nervous system activity (Parasuraman & Rizzo, 2007) with changes occurring in cardiac activity, brain activity, respiratory activity, speech measures, and eye activity. Other studies also reveal that cognitive demands have direct correlation with changes in both theta and alpha frequencies in the brain (Li, Dai, and Chang, 2025). Thus, by measuring autonomic system activity or brain activity, the mental state of an individual can be estimated (Miller, 2001).

EEG and Mental State-Classification

The electroencephalography (EEG) has been identified as an effective way of tracking brain activity during cognitive stress, meditation, and relaxation; making it ideal for evaluating mental states (Kotte et al., 2025). By measuring changes in both theta (4–7 Hz) and alpha (8–13 Hz) power (Li, Dai, and Chang, 2025), an increase in frontal theta power and a corresponding parietal alpha power decrease would indicate increasing cognitive workload (Popov et al, 2018). While EEG can track brain activity, utilizing machine learning techniques can help in examining, classifying, and processing EEG signals to accurately understand brain activity (Ahire et al., 2023). Despite the sensitivity of EEG, it has been identified to be susceptible to motion artifacts, muscles activity, and environmental noise (Stalin et al., 2021), which can negatively impact its application in operational environments (Alreshidi et al., 2023).

Peripheral Physiological Measurements

Electrocardiography (ECG), respiration, and galvanic skin response (GSR) represent some of the important peripheral measures used for cognitive-state assessment. ECG derivatives such as heart rate (HR) and heart rate variability (HRV) are some of the most widely used physiological markers of mental workload. According to Shaffer and Ginsberg (2017), an increase in cognitive effort and stress leads to a reduction in HRV, especially in the high-frequency bands. However, an increase in HR is an indication of increasing cognitive load (Wittels et al., 2024). Respiration rate and variability provide additional information about autonomic responses to cognitive demand. According to Grassmann et al. (2016), elevated workload is typically associated with faster, shallower breathing patterns and reduced respiratory variability. The Galvanic Skin Response (GSR) describes the electrodermal activity, also known as sweat gland permeability. It is a non-invasive method of capturing autonomic nerve response (Goshvarpour, Abbasi, and Goshvarpour, 2014) by measuring skin resistance to small electrical currents. Studies have shown

that an increase in cognitive workload leads to an increase in GSR; providing a basis for discriminating cognitive load using machine-learning algorithms (Nature Publisher Group, 2025). GSR supports the determination of stress and complements the detective capability of the ECG and respiration data (Crijnen, 2019). Including these measures in mental state studies would provide a more comprehensive picture of operator state and improve the diagnostic process (Grassmann et al., 2016).

Machine-Learning for Mental State-Classification

Recent studies in the field of neuroergonomics have seen the integration of Machine Learning techniques with attentional states studies for the continuous assessment of human functional state (Elena and Anastasia, 2026). The integration has contributed significantly to the advancement of cognitive workload detection by making real-time predictions of cognitive depletion; providing insights into mental and temporal demands of operators (Conklin et al., 2026). The machine learning models are able to automatically learn hierarchical representations from raw data, and also model nonlinear interactions across physiological features using various techniques and algorithms (He et al., 2025; Das Chakladar and Roy, 2024).

METHODOLOGY

In this study a comparative machine learning framework was employed to evaluate the differential sensitivity of electroencephalography (EEG), electrocardiography (ECG), respiration (RESP), and galvanic skin response (GSR) for classifying aviation-related mental states. The analysis was designed to address two complementary questions: first, how effectively each physiological modality independently discriminated among the mental states; and second, whether combining physiological modalities improved classification relative to single-modality models.

Data Acquisition

A 20-electrode EEG system in the standard 10–20 format + POz (Fz, Cz, Pz, F3, F4, C3, C4, P3, P4, O1, O2, T5, T3, F7, Fp1, Fp2, F8, T4, and T6 with Linked Mastoids) with sampling rate of 256 Hz was used. The captured signals involved a motion-based flight simulator, and another involved a non-flying environment (Alreshidi, Moulitas and Jenkins, 2023). The experimental design involved a pilot exposed to situations that elicited Channelized Attention, Diverted Attention (DA), and Startle/Surprise (SS) in a non-flight environment. The pilot also did a complete flying simulation using the motion-based flight simulator for take-off, flight, and landing.

Feature Extraction

Physiological recordings corresponding to Channelized Attention (CA), Diverted Attention (DA), Startle/Surprise (SS), and Line Oriented Flight

Training (LOFT) were organized and assigned their respective class labels for feature extraction. Features were also concentrated to create peripheral feature sets by combining ECG, respiration, and GSR; and combining EEG, ECG, respiration, and GSR to form the full multimodal fusion.

Classification Models

In this study, four supervised machine learning algorithms were evaluated: Logistic Regression, Support Vector Machine (SVM), Random Forest, and Multilayer Perceptron. The four algorithms were applied consistently across the physiological feature configurations to determine which of the algorithms worked best for each modality. Classification performance was evaluated using the Macro-F1 score and recall.

RESULTS AND DISCUSSION

The primary objective of this study is to investigate the differential sensitivity of EEG, ECG, Respiration, and Galvanic Skin Response (GSR) for mental-state classification. The secondary objective is to study whether multimodal physiological fusion improves classification performance compared to individual physiological modalities. From the data gathered, four mental states comprising of Channelized Attention (CA), Diverted Attention (DA), Startle/Surprise (SS), and Line-Oriented Flight Training (LOFT) were classified using four machine learning algorithms to determine the best algorithm for each modality.

From Table 1, the multimodal physiological fusion categorized as ALL consistently outperformed all the individual physiological modalities by producing the highest classification performance across all classifiers. Logistic Regression achieved the best overall results (Macro-F1 = 0.973), followed closely by the SVM (Macro-F1 = 0.963) and MLP (Macro-F1 = 0.956). These values substantially exceeded the performance of every individual physiological modality. This suggests that mental state transitions are multidimensional involving simultaneous changes in cortical, cardiovascular, respiratory, and electrodermal systems (Parasuraman and Wilson, 2008; Fairclough, 2009) requiring complementary information from multiple sources.

Secondly, EEG emerged as the most discriminative single modality; producing the highest classification performance regardless of the machine learning algorithm used. MLP achieved the highest EEG performance (Macro-F1 = 0.930). Following closely are Logistic Regression (0.921), SVM (0.874), and Random Forest (0.834). This positions the EEG as one of the most sensitive physiological indicators of attentional state (Borghini et al., 2014; Roy et al., 2016); due to its ability to directly measure cortical electrical activity associated with executive control, working memory, attention allocation, and decision making (Logah et al, 2025).

Thirdly, the peripheral (ECG + Respiration + GSR) physiological fusion became substantially more informative when integrated than when analyzed independently. The fusion substantially improved classification accuracy

with a maximum Macro-F1 score of 0.902 using the SVM classifier. This represents an approximately 18–37% point increment over the individual peripheral modalities. Out of the individual peripheral modalities, ECG produced the highest classification performance with Random Forest classifier achieving the best ECG result (Macro-F1 = 0.721). These findings suggest that cardiovascular dynamics contain meaningful information regarding mental-state transitions, although their discriminative capability remains substantially lower than EEG. Heart rate and heart-rate variability have long been recognized as indicators of sympathetic-parasympathetic balance during cognitively demanding tasks (Taelman et al., 2011). Increasing cognitive workload has been identified to be linked with increasing heart rate and decreasing heart rate variability (Kelkar et al., 2024; Logah et al., 2025). However, cardiovascular responses are influenced by numerous non-cognitive factors such as respiration, physical movement, and emotional state, thereby limiting their specificity for cognitive-state discrimination.

Respiration on the other hand produced moderate classification performance with SVM producing the highest Macro-F1 score of 0.633. Respiratory responses generally occur more slowly than cortical activity; therefore, may be less sensitive to rapid transitions between attentional states (Grassmann et al., 2016). In view of this, respiration alone may not provide sufficient temporal resolution for accurately discriminating complex mental states. Similarly, GSR exhibited moderate performance with SVM achieving the highest Macro-F1 score of 0.633. Primarily, electrodermal activity reflects sympathetic nervous system activation and therefore responds strongly to emotional arousal, stress, and surprise (Boucsein, 2012). This physiological characteristic could probably explain why GSR contributes useful information for the detection of highly aroused conditions such as the startle/surprise state but exhibits reduced sensitivity for distinguishing between more cognitively oriented conditions such as channelized attention and diverted attention.

Table 1: Model summary.

Modality	Model	Macro_F1_Mean
ALL	Logistic Regression	0.973232323
ALL	MLP	0.956449106
ALL	Random Forest	0.911963282
ALL	SVM	0.9625
ECG	Logistic Regression	0.627302665
ECG	MLP	0.642505795
ECG	Random Forest	0.720891608
ECG	SVM	0.568767507
EEG	Logistic Regression	0.921095016
EEG	MLP	0.930157065
EEG	Random Forest	0.833832556

(Continued)

Table 1: Continued.

Modality	Model	Macro_F1_Mean
EEG	SVM	0.873737096
GSR	Logistic Regression	0.502507493
GSR	MLP	0.580412475
GSR	Random Forest	0.632766955
GSR	SVM	0.516160359
PERIPHERAL	Logistic Regression	0.842179732
PERIPHERAL	MLP	0.885536963
PERIPHERAL	Random Forest	0.84527306
PERIPHERAL	SVM	0.901961372
RESP	Logistic Regression	0.606682345
RESP	MLP	0.555872045
RESP	Random Forest	0.585986097
RESP	SVM	0.633033617

Fourthly, the study revealed interesting differences among the different machine learning algorithms. From Figure 1, Logistic Regression consistently achieved the highest performance for the multimodal (ALL) dataset, whereas the MLP performed best for EEG, Random Forest performed best for ECG and GSR, and the SVM achieved the highest performance for both respiration and peripheral (ECG + Respiration + GSR) physiological fusion. This suggests that classifier selection should be matched to the statistical characteristics of physiological features. For instance, EEG features likely show nonlinear interactions among spectral bands, thus, would benefit from neural-network representations, whereas tree-based methods more effectively capture nonlinear threshold relationships present in ECG and GSR features. The superior performance of the SVM on peripheral physiological fusion further suggests that kernel-based classifiers are particularly effective for heterogeneous physiological datasets.

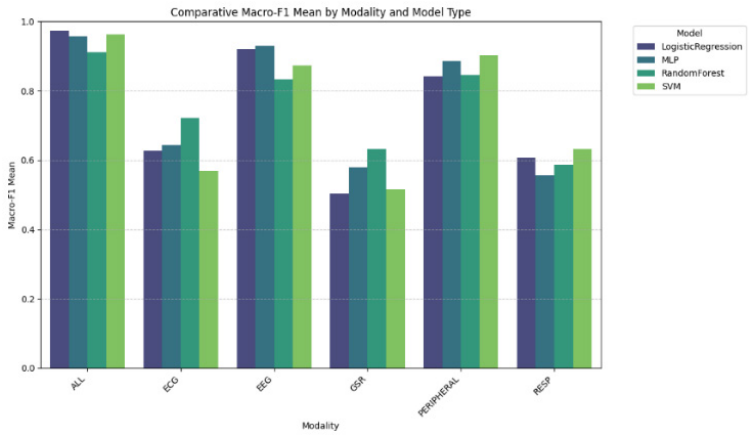


Figure 1: Comparative macro F1 mean by modality type.

Table 2 represents the modality sensitivity ranking obtained from the comparative machine learning analysis based on the Macro-F1 score. This primarily provides a balanced assessment of multiclass classification performance by equally weighting the precision and recall of each mental-state class (Chicco and Jurman, 2020). The results as shown in the table demonstrate a clear hierarchy in the discriminative capabilities of the physiological modalities (ALL>Peripheral>EEG>Respiration>ECG>GSR). The ranking provides important insight into the relative contributions of cortical and autonomic physiological signals for cognitive state recognition.

Among the modalities, multimodal fusion (ALL) achieved the highest overall classification performance with a Macro-F1 score of 0.963. This suggests that mental states cannot be fully characterized by utilizing a single physiological system. Mental-state transitions simultaneously influence cortical neural processing, cardiovascular regulation, respiratory activity, and electrodermal responses. Consequently, integrating these complementary physiological responses enables substantially richer characterization of operator mental state than any individual sensor alone (Fairclough, 2009; Parasuraman & Wilson, 2008). Peripheral physiological fusion follows as the second-highest ranked modality with a Macro-F1 score of 0.902. Even though the individual non-EEG modalities did not independently command impressive Macro-F1 scores, their fusion achieved significant performance.

Among the individual physiological modalities, EEG ranked highest, achieving a Macro-F1 score of 0.874. This was followed by respiration with a Macro-F1 score of 0.633, ECG (Macro-F1 score of 0.568), and GSR (Macro-F1 score of 0.516) in that order. This supports the assertion that EEG is the most informative single modality physiological measure of cognitive-state assessment (Borghini et al., 2014; Roy et al., 2016).

Table 2: Modal sensitivity ranking.

Modality	Macro_F1_Mean	Macro_Recall_Mean
ALL	0.9625	0.963333333
PERIPHERAL	0.901961372	0.903333333
EEG	0.873737096	0.875
RESP	0.633033617	0.655
ECG	0.568767507	0.58
GSR	0.516160359	0.543333333

Table 3 is the state sensitivity for the attentional states based on the six modalities used with class specific recall as the metric. From the table, the full multimodal fusion produced the strongest and most balanced state sensitivity. It correctly identified all DA and LOFT observations with recall values of 1 for both states. In the cases of CA and SS, it recorded recall values of 0.967 and 0.887 respectively. EEG achieved the

highest consistent sensitivity across the four states as far as the individual modalities are concerned. It reached recall values of 0.900 for CA, 0.853 for DA, 0.927 for SS, and 0.820 for LOFT. Its highest recall was observed in SS (0.927) which is characterized by rapid changes in attentional orientation, information processing, and cognitive control. ECG showed the most uneven pattern; reflecting a highly state-dependent sensitivity. It achieved a perfect recall for DA but only 0.487 for CA, 0.593 for SS, and 0.240 for LOFT. GSR discriminated DA and LOFT with recall values of 0.867 for both states but returned a recall value of 0.307 and 0.133 for both CA and SS respectively. Respiration was sensitive to CA (recall = 0.833) and LOFT (recall = 0.887) but weak for DA (recall = 0.253). The peripheral fusion achieved recall values of 0.893 for CA, 0.967 for DA, 0.787 for SS, and 0.967 for LOFT which indicate a balanced state sensitivity. Though the peripheral fusion results are similar to the EEG results in terms of balance state sensitivity, the peripheral fusion achieved higher sensitivity for DA and LOFT but lower sensitivity for CA and SS. This contrast suggests that the two physiological configurations captured different dimensions of the experimental states. EEG appeared particularly informative for cortical distinctions associated with CA and SS, whereas the fused autonomic features were especially effective for DA and the extended LOFT condition.

Table 3: State sensitivity.

Modality	CA_Recall	DA_Recall	SS_Recall	LOFT_Recall
ALL	0.96666667	1	0.8866667	1
ECG	0.48666667	1	0.5933333	0.24
EEG	0.9	0.85333333	0.9266667	0.82
GSR	0.30666667	0.86666667	0.1333333	0.866666667
PERIPHERAL	0.89333333	0.96666667	0.7866667	0.966666667
RESP	0.83333333	0.25333333	0.6466667	0.886666667

Figure 2 below represents the confusion matrix for the full multimodal model (ALL). In this model, EEG, ECG, respiration, and GSR features were combined for classification of Channelized Attention (CA), Diverted Attention (DA), Startle/Surprise (SS), and LOFT. The full model demonstrated a strong classification performance with only four misclassifications across the four attentional states. Apart from CA and SS that had a recall of 0.97 and 0.90 respectively, both DA and LOFT received a perfect recall. The misclassifications occurred between CA and SS, suggesting that there may be partial overlap between the physiological characteristics of CA and SS. The absence of confusion between DA and LOFT suggests that these states generated distinct multimodal physiological profiles within the available dataset. The results lend credence to the fact that a fusion of the modalities provide better results compared to individual modalities.

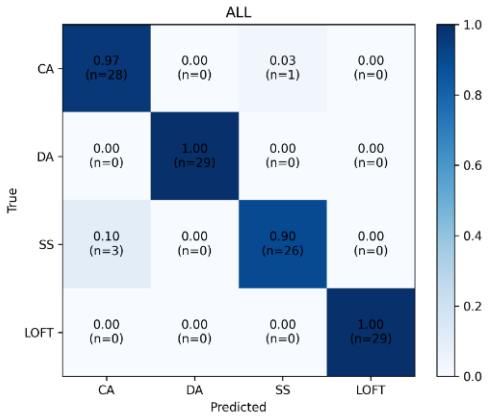


Figure 2: Multimodal confusion matrix.

CONCLUSION

The study conducted a comparative machine learning analysis of the differential sensitivity of electroencephalography (EEG), electrocardiography (ECG), respiration, and galvanic skin response (GSR) for distinguishing between four attentional states. Four different machine learning algorithms: Logistic Regression, Support Vector Machine (SVM), Random Forest, and Multilayer Perceptron (MLP) were evaluated using the individual modalities, peripheral physiological fusion (ECG + respiration + GSR), and a full multimodal fusion (EEG + ECG + respiration + GSR). From the analysis, it was observed that attentional state classification is strongly influenced by the physiological information available to the classifier. Even though EEG provides the strongest and most consistent discriminatory capability due to its ability to directly measure cortical electrical activity, integrating complementary physiological systems provide more complete representation of the operator state than using a single modality. Another important finding concerns the relative importance of physiological modality selection versus machine-learning algorithm selection. Although there were performance differences among the different machine learning algorithms, the differences are less compared with the differences observed in the different modalities. This suggests that even though the selection of an appropriate machine learning approach is important, choosing the right physiologically informative sensor may produce greater gains in neuroergonomic studies.

LIMITATIONS

The study focused on data collected from one participant, limiting the generalizability to other pilots. Future research should therefore replicate the analysis using a larger participant size before the results can be generalized to other pilots or operational environments.

SOURCE OF DATA

https://dspace.lib.cranfield.ac.uk/handle/1826/21792

REFERENCES

Ahire, N., Awale, R.N., Patnaik, S. & Wagh, A. (2023). "A comprehensive review of machine learning approaches for dyslexia diagnosis", Multimedia Tools and Applications, vol. 82, no. 9, pp. 13557–13577.

Alreshidi, I., Moulitsas, I. & Jenkins, K.W. (2023). "Multimodal Approach for Pilot Mental State Detection Based on EEG", Sensors, vol. 23, no. 17, pp. 7350.

Alshehhi, A., Abbasi, Q.H. & Ghannam, R. (2026) "Photoplethysmography for measuring cognitive load in XR environments: A systematic review", Advanced Sensor Research, vol. 5, no. 3.

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

Boucsein, W. (2012) Electrodermal activity. 2nd edn. Springer.

Chicco, D. and Jurman, G. (2020). "The advantages of the Matthews correlation coefficient over F1 score and accuracy in binary classification evaluation", BMC Genomics, vol. 21, no. 1, article 6.

Conklin, S., Kargosha, G., Tu, J., Bansal, K., Dang, Q., Brooks, J. & Kucukosmanoglu, M. (2026). "Cognitive biotypes identified through ECG-derived workload and behavioral accuracy", Scientific Reports (Nature Publisher Group), vol. 16, no. 1, p. 9934.

Crijnen, J.A. (2019). Predicting a pilot's cognitive state from physiological measurements. Tilburg University.

Das Chakladar, D. & Roy, P.P. (2024). "Cognitive workload estimation using physiological measures: A review", Cognitive Neurodynamics, vol. 18, no. 4, pp. 1445–1465.

Dias, R.D., Michael, S.H., Rayan, H., Yule, S., McGinnis, L., Skipworth, R.J.E., Mann-Salinas, E. & Davis, W.T. (2026). "Heart rate variability and body motion as digital biomarkers of task workload during military en route critical care simulations", Sensors, vol. 26, no. 11, article 3596.

Elena, K. & Anastasia, K. (2026). "Machine learning classification of physiological dynamics during standardized task-demand transitions", Bioengineering, vol. 13, no. 8, article 898.

Fairclough, S.H. (2009). "Fundamentals of physiological computing", Interacting with Computers, 21(1–2), pp. 133–145.

Goshvarpour, A., Abbasi, A. & Goshvarpour, A. (2014). "Impact of music on college students: Analysis of galvanic skin responses", Applied Medical Informatics, vol. 35, no. 4, pp. 11–20.

Grassmann, M., Vlemincx, E., von Leupoldt, A., Mittelstädt, J.M. & Van den Bergh, O. (2016). "Respiratory changes in response to cognitive load: A systematic review", Neural Plasticity, article 8146809.

Grassmann, M., Vlemincx, E., von Leupoldt, A. & Van den Bergh, O. (2016). "The role of respiratory measures to assess mental load in pilot selection", Ergonomics, vol. 59, no. 6, pp. 745–753.

He, M., Abdulrahman, A., Clore, J.L., Sowers, R.B., Hsiao-Wecksler, E. & Hernandez, M.E. (2025). "Scoping review of ML approaches in anxiety detection from in-lab to in-the-wild", Applied Sciences, vol. 15, no. 18, article 10099.

Kelkar, A., Natarajan, S., Kothari, A. & Bolisetty, M. (2024). "Comparison of cognitive workload and surgical outcomes between a three-dimensional and conventional microscope macular hole surgery", BMC Ophthalmology, vol. 24, pp. 1–7.

Kotte, S., Elkhouly, A., Abd Malek, M.F. & Abohaia, Z. (2025). "Deep learning-based EEG mental state classification to support mental focus in female cricketers", Discover Artificial Intelligence, vol. 5, no. 1, article 350.

Logah, F.X., Seong, Y., Baanye, J., Wilson, M., Adamolekun, A., Yi, S. & Kwakye, K. (2025). "A review of the impact of cognitive workload on reaction time and performance", SSRN.

Miller, S. (2001). "Workload measures. National Advanced Driving Simulator", Iowa City, United States.

Parasuraman, R. & Rizzo, M. (2007). "Introduction to neuroergonomics", in Neuroergonomics: The Brain at Work, pp. 3–12.

Parasuraman, R. and Wilson, G.F. (2008). "Putting the brain to work: Neuroergonomics past, present, and future", Human Factors, vol. 50, no. 3, pp. 468–474.

Popov, T., Popova, P., Harkotte, M., Awiszus, B., Rockstroh, B. & Miller, G.A. (2018). "Cross-frequency interactions between frontal theta and posterior alpha control mechanisms foster working memory", NeuroImage, vol. 181, pp. 728–733.

Roy, R.N., Bonnet, S., Charbonnier, S. & Campagne, A. (2016). "Mental fatigue and working memory load estimation: Interaction and implications for EEG-based passive BCI", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 24, no. 12, pp. 1224–1234.

Shaffer, F. & Ginsberg, J.P. (2017). "An overview of heart rate variability metrics and norms", Frontiers in Public Health, vol. 5, article 258.

Stalin, S., Roy, V., Shukla, P.K., Zaguia, A., Khan, M.M., Shukla, P.K. & Jain, A. (2021). "A machine learning-based big EEG data artifact detection and wavelet-based removal: An empirical approach", Mathematical Problems in Engineering.

Taelman, J., Vandeput, S., Spaepen, A. & Van Huffel, S. (2011). "Influence of mental stress on heart rate and heart rate variability", in Proceedings of the 4th European Conference of the International Federation for Medical and Biological Engineering, pp. 1366–1369.

Tierney, M., Adeyemi, E.O. & Bommer, S. (2025). "A systematic literature review: Cognitive workload assessment in human factors research", Advances in Human-Computer Interaction.

Wittels, H.L., S, H.W., Wishon, M.J., Vogl, J., Onge, P.S., McDonald, S.M. & Temme, L.A. (2024). "Examining the influence of cognitive load and environmental conditions on autonomic nervous system response in military aircrew: A hypoxia–normoxia study", Biology, vol. 13, no. 5, article 343.