International Journal of Industrial Engineering, 32(6), 1646-1660, 2025

ELECTROENCEPHALOGRAPHY (EEG) SIGNAL-BASED MENTAL WORKLOAD DETECTION FOR INTELLIGENT WORKSTATION MACHINE COGNITIVE TASK

Xiaoli Wu[1,2,3], Biao Yan[1,2,3*], Weiyi Han[1,2,3], Lan Zhang[1,2,3], Xiao Liu[1,2,3], Yuhan Wu[1,2,3], Xinyue Zhang[1,2,3], Zhiyu Wang[1,2,3], and Yuhang Zou[1,2,3]

[1]School of Design Art and Media
Nanjing University of Science and Technology
Nanjing, China
[*]Corresponding author's e-mail: yanbiao0109@163.com

[2]Research Centre of Human Cyber Physical Integration and Intelligent Interaction
Nanjing University of Science and Technology
Nanjing, China

[3]Key Laboratory of Ministry Industry and Information Technology for Language Information Processing and Application
Nanjing, China

The intelligent workstation serves as a cornerstone in the advancement of intelligent and digital manufacturing systems, with operators' roles increasingly shifting from physical labor to cognitive effort. To examine the effect of cognitive task difficulty on mental workload (MWL), operational tasks of low, medium, and high difficulty were designed as experimental conditions. The National Aeronautics and Space Administration Task Load Index (NASA-TLX), task performance metrics, and electroencephalographic (EEG) data were utilized to compare differences across the three levels. The results showed that, under high task difficulty, the average whole-brain power spectral density (PSD) in the Theta, Alpha, and High Beta frequency bands was significantly elevated among the 20 participants. Further electrode-level analysis revealed that the sensitivity and discriminative capacity of EEG signals varied across frequency bands and electrode sites. These EEG-based indicators demonstrate strong potential as neurophysiological biomarkers for differentiating levels of MWL and detecting operator overload in intelligent workstation contexts.

Keywords: Intelligent Workstation Machine, MWL, EEG, PSD.

(Received on March 31, 2025; Accepted on September 21, 2025)

1. INTRODUCTION

Industry 4.0 is increasingly characterized by large-scale automation, ubiquitous connectivity, and artificial intelligence-driven technologies, leading to significantly more efficient manufacturing processes (Kadir *et al*., 2021; Passalacqua *et al*., 2025). Intelligent, complex, digitalized, systematic, and automated human-machine interaction (HMI) systems are progressively replacing traditional manual labor, marking a new era in manufacturing. The integration of intelligent technologies profoundly reshapes the relationship between operators' MWL and performance (Guo *et al*., 2021). Lodgaard *et al*. (2020) highlight that recognition, perception, and visualization technologies are central to smart manufacturing HMI interfaces, with real-time, automated analysis of production data—presented through intuitive visualizations—forming the foundation for informed decision-making in industrial operations. Amidst this shift toward greater customization, the intelligent workstation has emerged as a critical tool for digital factories, facilitating planning, accelerating workflow, and achieving workforce reduction alongside efficiency improvements. However, as operational demands transition from physical labor to cognitive effort, some operators may struggle to meet the escalating cognitive requirements (Demerouti, 2022). Consequently, monitoring operators' MWL provides a set of sensitive evaluation metrics that can serve as a scientific basis for task difficulty assessment and task allocation optimization.

As a multidimensional construct, MWL is influenced by task-related factors (e.g., demands, performance), operator characteristics (e.g., skill level, attention), and, to some extent, the environmental context in which the task is performed (Young *et al*., 2015). At present, MWL assessment methods are generally categorized into three types: performance-based, physiological, and subjective measurements. Performance-based measurement assesses MWL by evaluating the operator's task performance outcomes. Physiological measurement involves monitoring human physiological indicators to reflect psychological load, offering real-time, continuous insights into the operator's cognitive state. Notably, physiological indicators are objective in nature and can help mitigate the accuracy limitations inherent in subjective assessments (Boff *et al*., 1986).

Previous studies have confirmed that EEG power across various frequency bands is sensitive to fluctuations in MWL. For instance, Zahabi *et al*. (2025) employed EEG and the NASA-TLX to examine the effects of task difficulty and repeated training on MWL in virtual reality (VR) environments. Their findings underscored the potential of EEG as a tool for real-time MWL monitoring during VR-based forklift inspection tasks. Similarly, Naseri *et al*. (2025) investigated the impact of different MWL levels on brain activity during driving and reported significant EEG differences between high and low cognitive workload conditions, particularly in the delta and theta bands. Qin and Bulbul (2023) assessed the usability of augmented reality (AR) head-mounted displays (HMDs) in construction assembly tasks, conducting real-time physiological evaluations of users' MWL. Their results demonstrated that EEG-based MWL assessment can effectively reflect real-time fluctuations in workers' cognitive states during assembly operations. In a related study, Adachi *et al*. (2024) explored both mental and physical loads contributing to driver fatigue under varying traffic conditions. Changes in EEG spectral responses, especially within the theta and alpha bands, were found to correlate with fatigue, depending on the traffic environment.

Additionally, Ji *et al*. (2023) found that pilots' MWL during flight tasks in aircraft spatial environments was most strongly associated with EEG activity in the regional beta bands. Although these studies provide valuable insights, their findings may not be directly generalizable to the context of smart manufacturing. Consequently, further research is necessary to investigate how MWL influences operators' physiological responses and task performance in industrial settings.

Subjective assessment methods primarily involve the use of scales or questionnaires, among which NASA-TLX, SWAT, and SMEQ are the most widely adopted. The NASA-TLX scale has demonstrated strong reliability and validity (Hutson *et al*., 2024), encompassing six key dimensions—such as mental demand, time pressure, and frustration—thus providing a comprehensive reflection of participants' subjective workload experience. For example, Jajo and Biondi (2023) employed the Detection Response Task (DRT) as an objective measure of cognitive load alongside the NASA-TLX for subjective assessment, finding no significant differences in cognitive workload between partially automated and manual modes in either measurement type. Arana *et al*. (2023) utilized hierarchical task analysis and error identification to perform a cognitive workload assessment, with NASA-TLX results serving to quantify participants' subjective MWL. Yeh and Chao (2024) examined the effects of screen brightness and color combinations on drivers' visual performance and fatigue, collecting Critical Flicker Fusion frequency (CFF), touchscreen deviation data, and NASA-TLX scores. Their findings revealed that ambient illumination significantly influenced touchscreen deviations, CFF, and subjective workload, while color combinations notably affected touchscreen deviations and NASA-TLX ratings. Javernik *et al*. (2023) emphasized the role of human workload in human-robot collaboration, employing NASA-TLX to assess workers' workload and highlighting the need for personalized guidelines that account for individual differences in abilities, skills, and personality traits to optimize overall well-being and productivity. Özdemir *et al*. (2023) analyzed data from real-time simulations and NASA-TLX surveys, identifying the number of aircraft, controller performance, and interruption duration as key factors impacting pseudo-pilots' overall workload. Overall, the NASA-TLX has been extensively applied for subjective workload assessment across various domains, exhibiting excellent usability. Within the context of smart manufacturing, it also provides valuable real-time evaluation, complementing physiological measurements and thereby enhancing the robustness and reliability of workload research.

In summary, existing studies have investigated the impact of intelligent information processing task demands on operator MWL through human factors experiments across various fields. However, empirical research remains relatively scarce within the context of intelligent manufacturing, especially regarding operational scenarios in smart workstations. To address this gap, the present study adopts machine operation tasks performed at smart workstations as the experimental paradigm to systematically examine the influence of task difficulty on operator MWL. By integrating the NASA-TLX subjective rating scale, performance metrics, and multimodal physiological data, a comprehensive assessment of MWL is provided. Notably, EEG was employed to analyze changes in PSD across different levels of task difficulty, aiming to identify neural features that are highly sensitive to MWL fluctuations. It is hypothesized that PSD values in the Theta and High Beta frequency bands will increase significantly under high workload conditions. Importantly, the extracted EEG frequency features demonstrate strong sensitivity and discriminative power within the unique human-machine interaction environment of smart workstations. These findings reveal both novelty and practical significance, offering a robust neurophysiological foundation for the development

of a quantifiable task difficulty classification system in future applications.

2. METHOD

2.1 Participants

A total of 22 mechanical engineering students from Nanjing University of Technology participated in the study, comprising 11 males and 11 females aged between 20 and 24 years. Inclusion criteria required participants to be right-handed, as left-handed individuals exhibit greater interhemispheric variability (left, right, or bilateral activation) during task performance, whereas right-handed individuals demonstrate more consistent activation patterns. Exclusion criteria included color blindness and color vision deficiency, which were screened through self-reporting by the participants. All participants had normal or corrected-to-normal vision. This study was approved by the Ethics Committee of Nanjing University of Science and Technology, and informed consent was obtained from all participants prior to the experiment.

2.2 Design

To investigate the correlation between operator MWL and EEG physiological indicators, this study employed a one-factor within-subjects experimental design. MWL was manipulated by varying the difficulty of tasks performed on an intelligent workstation, consisting of three levels: low difficulty (C1), medium difficulty (C2), and high difficulty (C3). As task difficulty increased, both the number of steps required for completion and the complexity of operations correspondingly increased.

2.3 Equipment

The experiment was programmed using E-Prime 3.0, a software platform designed to present interactive task interfaces while supporting operator interaction via mouse input. Its primary function in this study was to record and log the timing and details of human-computer interaction behaviors, including mouse clicks, keyboard inputs, and other operational actions. The experimental hardware system consisted of multiple display screens, control modules, servers, and associated equipment.

EEG signals were simultaneously recorded from 32 electrode sites using the Brain Products EEG acquisition system (Brain Products GmbH, Germany). The electrode locations included: FP1, F3, F7, FT9, FC5, FC1, C3, T7, TP9, CP5, PZ, P3, P7, O1, OZ, O2, P4, P8, TP10, CP6, CP2, CZ, C4, T8, FT10, FC6, FC2, F8, F4, FP2, and FPZ. The data were recorded with a bandwidth of 0–100 Hz and a sampling rate of 1000 Hz. The FZ electrode was used as the reference. The experimental setup used for EEG data acquisition is shown in Figure 1.

Figure 1. Configuration of the EEG acquisition system. (L: Experimental equipment R: Laboratory environment)

2.4 Stimulus material

The experimental materials consisted of the operational task interface from a smart factory workstation machine system, as illustrated in Figure 2. The interface was configured at a resolution of 1440 × 1080 pixels, in BMP format with an 8-bit color depth, to ensure compatibility with the visual presentation requirements of EEG experiments. The task difficulty was manipulated as the experimental variable and categorized into three levels: low, medium, and high. As task difficulty

increased, the number of operational steps required also increased. All three task levels were executed within the same operating system environment.

2.5 Task and procedure

The experiment required participants to complete a full task workflow consisting of three subtasks with varying levels of difficulty: a low-difficulty input task, a medium-difficulty completion task, and a high-difficulty monitoring task (see Table 1). These were designated as C1 (input task phase), C2 (completion task phase), and C3 (monitoring task phase), respectively. During phases C2 and C3, participants were also required to respond to alarms by providing keystroke feedback. Three types of alarms were included: low dosage alarm, product defect alarm, and machine fault alarm. In phase C2, a single visual alarm was presented, while in phase C3, two simultaneous visual alarms were displayed, thereby increasing the mental workload.

Table 1. Difficulty classification of intelligent workstation machine operation tasks

	C1	C2	C3
Main tasks	Input phase: change the machine status to running, modify staff	Complete phase: registering the number of products produced and modifying the machine to a shutdown state	Monitoring phase: modification of task orders, memory yield, memory movement rate, registration of defective products
Sub-tasks	--	Keystroke feedback to complete visual alerts	Keystroke feedback to complete visual alerts

The experiment comprised two phases: a pre-test phase and a formal test phase. During the pre-test, participants signed an informed consent form and became familiar with the experimental tasks through a simulation based on detailed task instructions. In the formal test phase, each participant wore an EEG cap and completed one of the three tasks corresponding to the designated difficulty level. After completing each task condition, participants took a short break and filled out a paper-based questionnaire assessing their perceived MWL, affective state, and memory of visual stimuli.

2.6 Measurement methods and metrics

The EEG indicator measures were selected from whole brain Delta, Theta, Alpha, Low beta, and High Beta, 5 different frequency bands of EEG signals of the whole brain average, with each electrode's power spectral density as an EEG evaluation index. The specific EEG signal frequency bands and the change of each frequency band with task difficulty are shown in Table 2.

Table 2. EEG signal frequency bands

Brain bands	**Frequency range (HZ)**	**State of mind**	**Variation of band power with increasing task demand**
Delta	0.5-4	Deep relaxation and restorative sleep	Increases with increasing task difficulty (Ruan *et al*., 2022)
Theta	4-8	Sleepiness, lethargy, contemplation, and dreaming	Increase, associated with fatigue and information retrieval (Borghini *et al*., 2012)
Alpha	8-13	Relaxed, calm, and clear-minded	Increases, as information processing for secondary tasks increases (Hebbar *et al*., 2021)
Low Beta	13-21	An alert, actively focused, busy, and anxious mind	Decreasing (Pavlov & Kotchoubey, 2017)
High Beta	21-30	Focused, quick thinking, working	Increasing (Pavlov & Kotchoubey, 2017)

3. RESULTS

Valid data were obtained from 20 participants, and statistical analyses were conducted using SPSS Statistics 26.0, with a significance level set at 0.05. To examine the main effects of task complexity factors on the measured indicators, a within-subject one-way repeated measures ANOVA and the Friedman test were employed, and the Bonferroni multiple comparison method was used for post hoc tests.

3.1 NASA-TLX scale scores

The six dimensions of the NASA-TLX scale were evaluated using unweighted raw scores, without employing a weighting procedure involving pairwise comparisons among the dimensions. The overall workload score was calculated as the arithmetic mean of the dimension scores. Previous research has shown that this simplified approach is widely adopted in experimental studies to streamline the assessment process (Biernacki & Lewkowicz, 2024). The results of the NASA-TLX for tasks with varying levels of intelligent workstation operation difficulty are presented in Figure 2. The scale scores exhibited an increasing trend as task difficulty increased. The NASA-TLX scores were subjected to the Shapiro-Wilk test for normality, revealing that the samples under conditions C1, C2, and C3 all followed a normal distribution ($P > 0.05$). Mauchly's test for sphericity indicated that the samples did not meet the sphericity assumption ($\chi^2(20) = 17.339$, $P < 0.05$). A one-way repeated-measures ANOVA with Greenhouse-Geisser correction revealed a significant main effect of operational task difficulty on NASA-TLX scale scores ($F(1.24, 23.48) = 42.73$, $P = 0.00 < 0.01$, $\eta_p^2 = 0.692$). The between-subjects effect test further revealed that NASA-TLX scale scores were significantly higher for task difficulty C2 (M=46.32, SD=10.93, 95%CI[41.21, 51.44]) than task difficulty C1 (M=39.38, SD=12.35, 95%CI[33.60, 45.15]), P<0.01; task difficulty C3 (M=47.55, SD=10.87, 95%CI[42.46, 52.64]) than task difficulty C2 (M=46.32, SD=10.93, 95%CI[41.21, 51.44]), p<0.05,; NASA-TLX scale scores for task difficulty C3 (M=47.55, SD=10.87) were significantly higher than task difficulty C1 (M=39.38, SD=12.35, 95%CI[33.60, 45.15]), P<0.01.

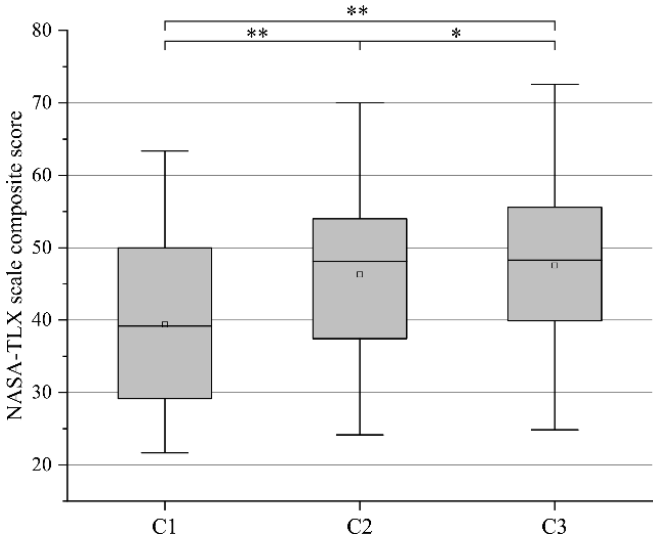


Figure 2. NASA-TLX scale scores at different cognitive task difficulties (*P<0.05, **P<0.01)

3.2 Task performance

Operating correctness was defined as the ratio of the number of tasks correctly completed by participants to the total number of task attempts. The results of operating correctness across different levels of intelligent workstation task difficulty are presented in Figure 3. A decreasing trend in operating performance was observed as task difficulty increased. The Shapiro–Wilk test indicated that the data did not follow a normal distribution ($P < 0.05$), the Friedman test revealed a significant main effect of operational task difficulty on the correctness of the operational task ($\chi^2(2) = 23.507$, $p<0.01$) with an effect size of $W_{Kendall}=0.588$. Bonferroni-adjusted post hoc tests revealed that the correctness rate of task difficulty C3 (M=64%, SD=17.89%, 95%CI[55.63, 72.37]) was significantly lower than that of task difficulty C1 (M=94%, SD=11.42 %,

95%CI[88.65, 99.35]), p<0.01; the correct rate for task difficulty C3 (M=64%, SD=17.89%, 95%CI[55.63, 72.37]) was significantly lower than that for task difficulty C2 (M=85.71%, SD=14.66%, 95%CI[78.85, 92.57]), p<0.01; the main effect of correct rate for task difficulty C2 was not significant with task difficulty C1, p=0.54.

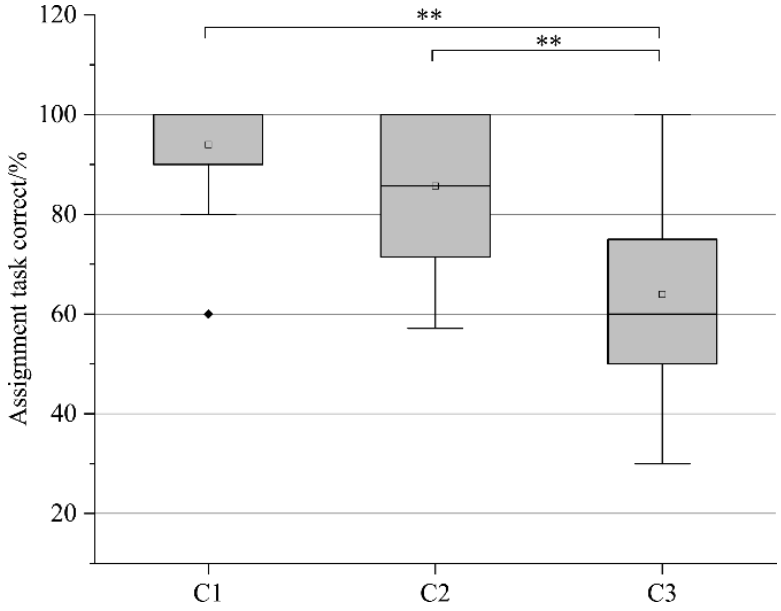


Figure 3. Correctness rate under different cognitive task difficulty conditions (*P<0.05, **P<0.01)

3.3 EEG

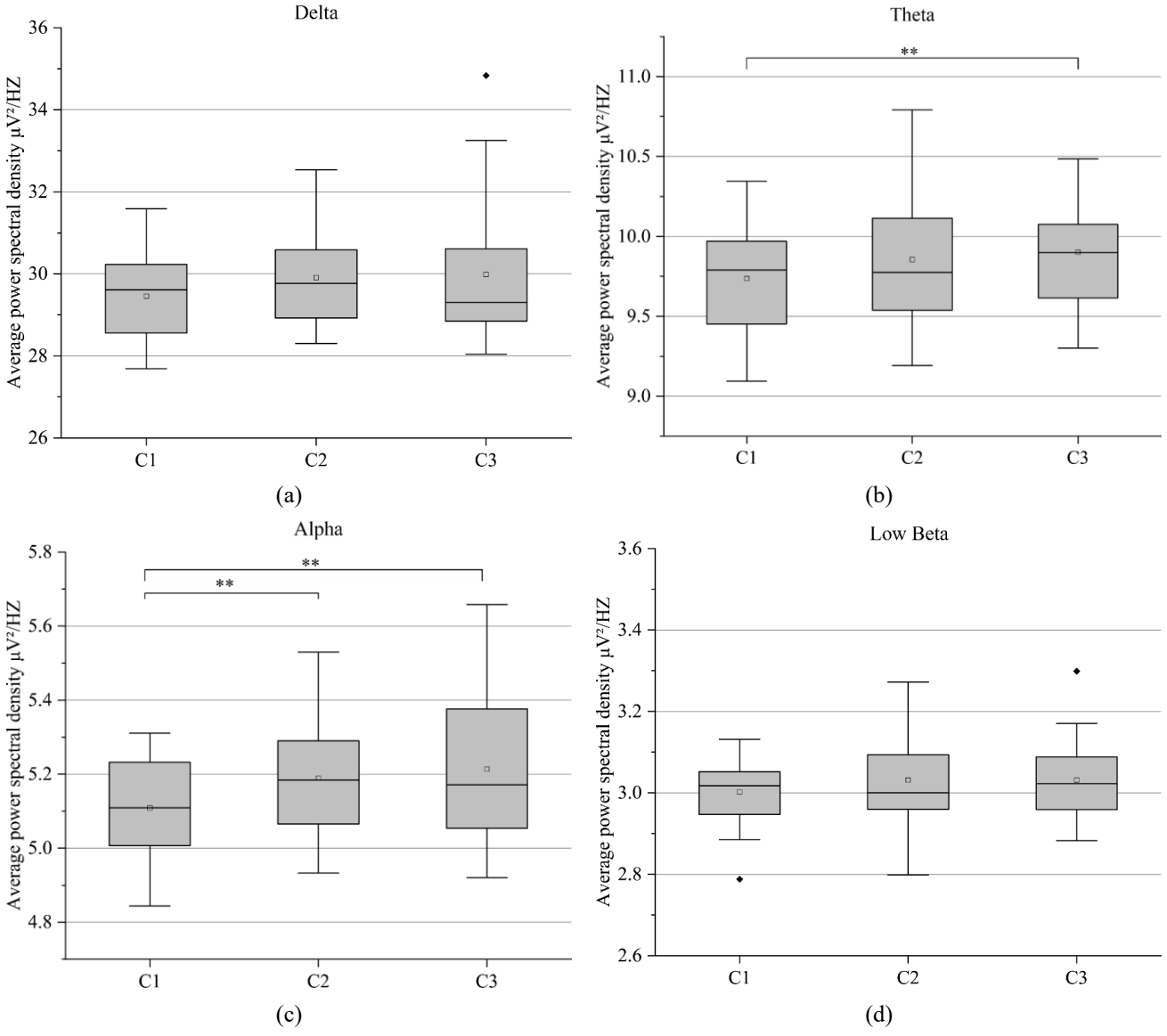
Raw EEG signals were preprocessed using MATLAB R2019 and the EEGLAB 14.1.1b toolbox. Electrode positioning was first conducted, and the common average reference was applied. A finite impulse response (FIR) band-pass filter with a frequency range of 0.5–50 Hz was used to filter the data. EEG signal segments collected 11 seconds before and after the alarm onset in the C1, C2, and C3 tasks were selected for further analysis. Independent component analysis (ICA) was then performed, and non-neural components—such as those related to ocular (EOG), muscular (EMG), and cardiac (ECG) artifacts—were identified and removed using EEGLAB plugins. For EEG analysis, power spectral density (PSD) was selected as the primary index. PSD was extracted from the preprocessed EEG data using MATLAB's periodogram function, and the average power spectrum was calculated for each task condition. Finally, the power values for five standard EEG frequency bands were derived from the spectral data: Delta (0.5–4 Hz), Theta (4–8 Hz), Alpha (8–13 Hz), Low Beta (13–21 Hz), and High Beta (21–30 Hz).

(1) Differences in Mean Absolute Power across the Brain

The Shapiro-Wilk test confirmed the normal distribution of PSD data for the Theta, Alpha, Low Beta, and High Beta frequency bands (all $P > 0.05$). As shown in Figure 4, repeated-measures ANOVA revealed statistically significant main effects of mental workload on whole-brain mean PSD in the Theta ($F = 4.641$, $P < 0.05$), Alpha ($F = 7.785$, $P < 0.01$), and High Beta ($F = 3.685$, $P < 0.05$) bands. Post-hoc tests further specified these effects: in the Theta band, PSD was significantly higher in C3 than in C1 ($P < 0.05$), with no significant differences between C2 and C1 or between C3 and C2. In the Alpha band, PSD in both C2 and C3 was significantly elevated compared with C1 ($P < 0.05$), whereas C3 and C2 did not differ significantly. Similarly, in the High Beta band, PSD in C3 was significantly greater than in C1 ($P < 0.05$), while neither C2 versus C1 nor C3 versus C2 reached statistical significance.

To provide a spatial visualization of the statistical results described above, Figure 5 displays the topographic distribution of whole-brain average absolute power across five frequency bands (Delta, Theta, Alpha, Low Beta, and High Beta) under three mental workload conditions (C1, C2, and C3). Overall, the topography reveals a clear macro-level pattern: as mental load increases, the scalp potential distribution shifts systematically from cool to warm color tones. This visual gradient corresponds to the statistically significant increases in mean power spectral density (PSD) shown in Figure 4, with the most pronounced changes observed in the Theta, Alpha, and High Beta bands.Specifically, under the C3 high-load condition, the

contraction of cool-colored areas and expansion of warm-colored regions across these bands spatially reflect enhanced cortical activation associated with high cognitive demand: the warm-area expansion in the Theta band is primarily concentrated in the frontal region (indicated by a white arrow); the reduction in activity and shift to warm tones in the Alpha band mainly occurs in the occipital region (indicated by an orange arrow); and the activation increase in the High Beta band is broadly distributed across the parietal region (indicated by a red arrow). Additionally, the expansion of warm areas in the Alpha band from C1 to C2—predominantly located in the occipital region (indicated by a purple arrow)—reflects significant topographic alteration associated with intermediate mental workload.



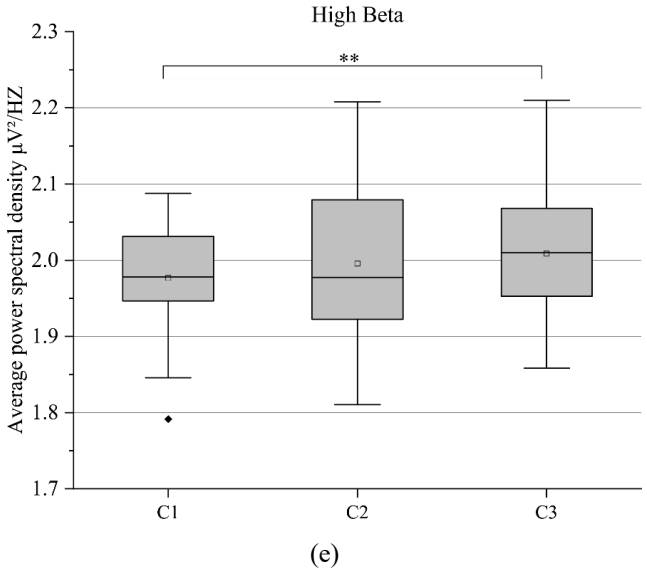


Figure 4. Average whole-brain PSD for each frequency band at different cognitive task difficulties (*P<0.05, **P<0.01)

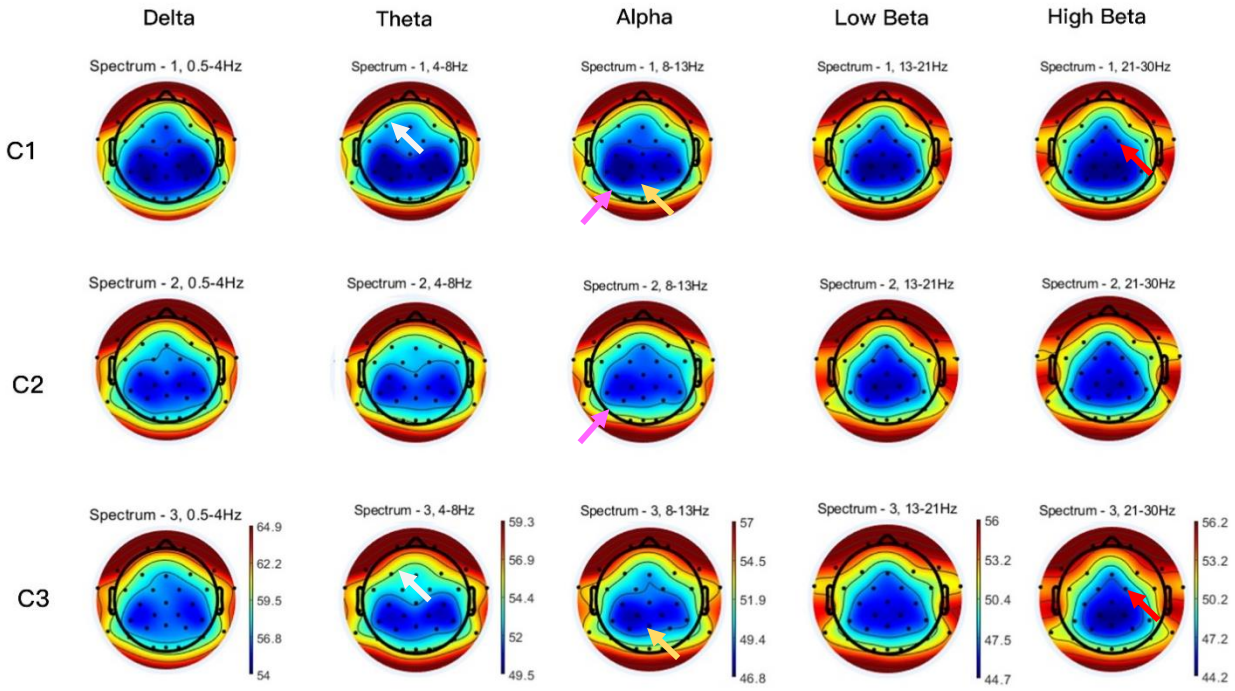


Figure 5. Topographic distributions of whole-brain average absolute power for five frequency bands under three mental workload conditions

(2) Variations in PSD Across Electrodes for Different MWL Tasks

Table 3 presents the results of repeated measures ANOVA or Friedman's test on the PSD at each electrode across different frequency bands under varying MWL task conditions. Only results with p-values less than 0.05 are shown. Although the Delta and Low Beta bands did not show significant differences at the whole-brain level, electrodes such as FP2, Oz, F7, and CP1 exhibited significant sensitivity to task difficulty within these bands ($P < 0.05$). Figure 6 illustrates the topographical distribution of electrodes showing significant PSD differences across frequency bands. Notable differences were observed in the Delta, Alpha, Low Beta, and High Beta bands within the right prefrontal and occipital regions ($P < 0.05$); the Theta band within the mid-frontal and parieto-occipital regions ($P < 0.05$); and the Alpha, Low Beta, and High Beta bands across bilateral frontal, parietal, and parieto-occipital regions ($P < 0.05$).

Table 4 presents the results of Bonferroni post-hoc tests on PSD at each electrode and frequency band under different MWL intelligent workstation tasks. The sensitivity of electrodes to task difficulty varied across frequency bands. Specifically, the FP2 electrode in the prefrontal region and the Oz electrode in the occipital region showed significant sensitivity in the Delta, Alpha, Low Beta, and High Beta bands ($p < 0.05$). The F7 electrode in the frontal region exhibited significant sensitivity in the Alpha, Low Beta, and High Beta bands ($p < 0.05$). Additionally, the FC5 electrode in the posterior frontal region and the P4 electrode in the parieto-occipital region showed significant sensitivity in the Alpha and High Beta bands ($p < 0.05$). Furthermore, significant differences in brain load changes between conditions C2 and C3 were found at the frontal electrodes FZ, F8, FC5, and CP5 within the Alpha and High Beta bands ($p < 0.05$), while no significant differences were observed at other frequency bands or electrodes.

Table 3. Repeated measures ANOVA or Friedman's test for PSD at each electrode in each frequency band under different cognitive task difficulty conditions

Electrode point	**Brain frequency band**				
	Delta	**Theta**	**Alpha**	**Low Beta**	**High Beta**
FP2	$\chi^2(2)=7.300$, $P=0.026^*$	-	$\chi^2(2)=6.100$, $P=0.047^*$	$\chi^2(2)=12.900$, $P=0.002^*$	$F=6506$, $P=0.04^*$
F3	-	$F=7.363$, $P=0.002^*$	-	-	-
FZ	-	-	$\chi^2(2)=6.300$, $P=0.043^*$	-	-
F8	-	-	-	-	$F=6.707$, $P=0.03^*$
F7	-	-	$F= 4.268$, $P=0.021^*$	$F=3.603$, $P=0.037^*$	$F= 3.406$, $P=0.044^*$
FC5	-	-	$\chi^2(2)=6.100$, $P=0.047^*$	-	$F= 4.125$, $P=0.024^*$
FC6	-	-	-	-	$F=7.091$, $P=0.02^*$
CP5	-	-	-	-	$F= 3.377$, $P=0.045^*$
CP1	-	$\chi^2(2)=7.500$, $P=0.024^*$	-	$\chi^2(2)=7.600$, $P=0.022^*$	-
P3	-	$\chi^2(2)=6.700$, $P=0.035^*$	-	-	-
P4	-	-	$\chi^2(2)=7.600$, $P=0.022^*$	-	$\chi^2(2)=6.100$, $P=0.047^*$
Oz	$F=4.225$, $P=0.022^*$	-	$\chi^2(2)=9.300$, $P=0.010^*$	$F=4.069$, $P=0.025^*$	$\chi^2(2)=7.900$, $P=0.019^*$

Note: * $P < 0.05$.

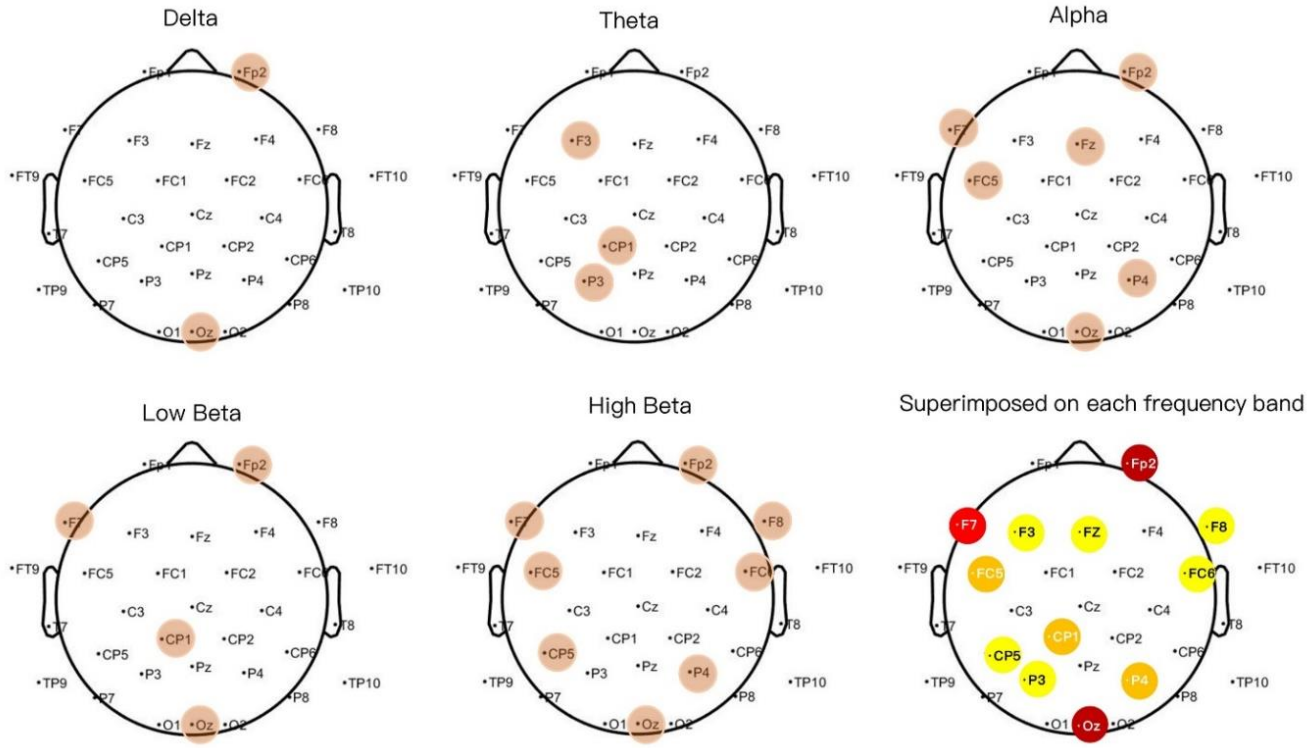


Figure 6. Location of electrode sites with significant PSD differences at different cognitive task difficulties. (indicates P<0.05)

Table 4. Bonferroni multiple comparison test of PSD across frequency bands for each electrode under different cognitive task difficulty conditions

Electrode point	Brain band and task difficulty										
	Delta	Theta		Alpha			Low Beta		High Beta		
	C1-C3	C1-C2	C1-C3	C1-C2	C1-C3	C2-C3	C1-C2	C1-C3	C1-C2	C1-C3	C2-C3
FP2	P= 0.007*	-	-	P= 0.040*	P= 0.027*	-	P= 0.004*	P= 0.001*	P= 0.042*	P= 0.00*	-
F3	-	-	P= 0.003*	-	-	-	-	-	-	-	-
FZ	-	-	-	-	-	P= 0.018*	-	-	-	-	-
F8	-	-	-	-	-	-	-	-	-	-	P= 0.004*
F7	-	-	-	-	P= 0.012*	-	-	P= 0.035*	-	P= 0.049*	-
FC5	-	-	-	-	P= 0.04*	P= 0.027*	-	-	p= 0.044*	-	-
FC6	-	-	-	-	-	-	-	-	-	P= 0.03*	-
CP5	-	-	-	-	-	-	-	-	-	P= 0.044*	P= 0.002*
CP1	-	P= 0.018*	P= 0.018*	-	-	-	P= 0.027*	P= 0.011*	-	-	-
P3	-	-	P= 0.011*	-	-	-	-	-	-	-	-
P4	-	-	-	P= 0.027*	P= 0.011*	-	-	-	P= 0.04*	P= 0.027*	-
Oz	P= 0.033*	-	-	P= 0.018*	P= 0.004*	-	P= 0.029*	P= 0.039*	P= 0.040*	P= 0.007*	-

Note: * P < 0.05.

4. DISCUSSION

4.1 NASA-TLX scale and performance analysis

In ergonomics research, the NASA-TLX scale and task performance are widely utilized as indicators of MWL, which is directly influenced by the consumption of cognitive resources and the degree of cognitive resource conflict (Vidulich & Tsang, 2012). The results of this experiment indicated that both NASA-TLX scores and task error rates increased significantly with rising task difficulty and the inclusion of additional subtasks. Notably, NASA-TLX scores were significantly higher for the medium- and high-difficulty tasks compared to the low-difficulty task. This may be attributed to the introduction of alerting tasks and elevated task demands, which reduced the operators' available attentional resources and subsequently led to increased subjective workload.

4.2 EEG power spectrum analysis

The results of the study indicated that as the difficulty of the operational task increased, participants' MWL levels also increased. Correspondingly, whole-brain average power in the Theta, Alpha, and High Beta frequency bands was enhanced. This finding is consistent with previous research: Theta-band EEG activity is positively correlated with cognitive demands and tends to increase when task difficulty rises or attention is heightened (Holm *et al*., 2009; Li *et al*., 2025). The Alpha band is known to be sensitive to information processing, particularly in secondary tasks, and its activity increases with greater cognitive engagement. High Beta power is strongly associated with cognitive load and tends to increase as workload intensifies (Pavlov & Kotchoubey, 2017; Yu *et al*., 2024).

To examine differences in brain activity under varying MWL conditions, one-way ANOVA and the Friedman test were conducted for each electrode (see Table 3). Theta activity was significantly elevated in the frontocentral and parieto-occipital regions; Low Beta activity was significantly increased in the frontal, parietal, and occipital regions; while Alpha and High Beta activity showed significant enhancement across widespread cortical areas.

As shown in Table 4, the FP2 electrode in the prefrontal cortex and the Oz electrode in the occipital region demonstrated high sensitivity to MWL differences in the Delta, Alpha, Low Beta, and High Beta bands, consistent with the findings of Dehais *et al*. (2019) and He *et al*. (2019). However, the degree of sensitivity varied across frequency bands. The F7 electrode in the frontal region also exhibited strong sensitivity to MWL differences in the Alpha, Low Beta, and High Beta bands, aligning with the results of Orun *et al*. (2019). Additionally, the FC5 electrode in the posterior frontal region and the P4 electrode in the parieto-occipital region showed high sensitivity to MWL differences in the Alpha and High Beta bands. Notably, although Delta and Low Beta frequencies did not show significant differences in the whole-brain mean power spectral density, the FP2, Oz, and F7 electrodes demonstrated diagnostic capability for distinguishing task difficulty. These findings offer supplementary physiological indicators for evaluating MWL under different cognitive task demands.

4.3 Limitations and future work

On one hand, it is crucial to shift research focus toward real-world workplace and industrial environments. However, the findings of this study were obtained in a controlled laboratory setting, underscoring the necessity to validate and replicate these results in representative, ecologically valid environments. On the other hand, individual differences in learning ability or personality traits may influence subjective perceptions of task difficulty (Wilson, 2002; Rozado & Dunser, 2015; Hu *et al*., 2022). Therefore, future research should extend beyond cross-sectional designs to incorporate longitudinal repeated-measures studies that account for intra-individual variability arising from interpersonal differences and contextual factors. Moreover, some scholars argue that individual physiological metrics may fail to reliably reflect task load changes due to measurement noise, artifacts caused by subject movement, and limited training data (Chanel *et al*., 2006; Wilson & Russell, 2007; Nilsson *et al*., 2022). By employing multiple measurement modalities concurrently, MWL assessment can achieve greater accuracy and stability (Hogervorst *et al*., 2014; Brouwer *et al*., 2017; Jiang *et al*., 2024). Thus, future studies aiming to differentiate MWL levels should incorporate a broader array of sensors. Moreover, in the current experiment, participants' visual conditions were assessed solely through self-report. Future studies might consider using standardized visual assessment tools to improve the objectivity of the findings.

With real-time EEG monitoring of MWL, work planners and manufacturing engineers can gain a deeper understanding of specific cognitive demands triggered by human-computer interaction tasks such as task complexity and task switching (Parmentier *et al*., 2020; Jin *et al*., 2024). Consequently, future smart manufacturing systems could dynamically adjust production procedures or operational instructions based on operators' real-time MWL levels. Similarly, the design and

presentation of interactive information within smart manufacturing environments can be optimized to help operators manage and reduce their mental load (Chen *et al*., 2016; Longo & Padovano, 2017; Landi *et al*., 2018; Chan *et al*., 2022). Therefore, optimizing MWL through integrated intelligent measurement systems can promote employees' mental health and enhance individual performance, enabling workers to better leverage the benefits brought by intelligent technologies. Conversely, excessive MWL in smart manufacturing contexts may increase the risk of operational errors or safety incidents, and negatively impact worker motivation as well as their physical and psychological well-being (Demerouyi, 2022).

5. CONCLUSION

This study, conducted on an intelligent workstation operation platform, comprehensively investigated operators' MWL states while performing tasks of varying difficulty levels. By designing tasks that elicit different MWL levels, EEG was utilized to analyze changes in PSD across multiple frequency bands during task execution. Additionally, a multidimensional assessment of MWL was performed by integrating subjective ratings from the NASA-TLX scale with objective task performance metrics, with the aim of identifying features highly sensitive to MWL fluctuations.

The results demonstrated that NASA-TLX scores, task accuracy rates, and average whole-brain PSD values in the Theta, Alpha, and High Beta frequency bands exhibited strong sensitivity to changes in MWL. Further analyses revealed significant differences in the sensitivity and discriminative capability of localized electrode signals across frequency bands: Delta activity was enhanced in the prefrontal and occipital regions; Theta activity was prominent in the fronto-central and parieto-occipital regions; Low Beta activity increased in the fronto-central, parietal, and occipital areas; while Alpha and High Beta activities showed widespread enhancement throughout the frontal, parietal, and occipital lobes. These findings provide crucial neurophysiological indicators for the real-time online monitoring of operator MWL. Moreover, the results offer empirical evidence to support the development of task difficulty classification systems and the optimization of task control processes in intelligent workstation cognitive operations, ultimately contributing to enhanced safety and efficiency in human-machine system performance.

REFERENCES

Kadir, B. A., & Broberg, O. (2021). Human-centered design of work systems in the transition to Industry 4.0. Applied ergonomics, 92, 103334.

Passalacqua, M., Pellerin, R., Magnani, F., Doyon-Poulin, P., Del-Aguila, L., Boasen, J., & Léger, P. M. (2025). Human-centered AI in industry 5.0: a systematic review. International Journal of Production Research, 63(7), 2638-2669.

Guo, S., Wanyan, X., Liu, S., Liang, C., & Chen, H. (2021). Influences of Intelligent design and information processing modality complexity on occupant mental workload. Acta Armamentarii, 42(2), 234.

Lodgaard, E., & Dransfeld, S. (2020). Organizational aspects for successful integration of human-machine interaction in the Industry 4.0 era. Procedia cirp, 88, 218-222.

Demerouti, E. (2022). Turn digitalization and automation into a job resource. Applied Psychology, 71(4), 1205-1209.

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: mental workload in ergonomics. Ergonomics, 58(1), 1-17.

Boff, K. R., Kaufman, L., & Thomas, J. P. (Eds.). (1986). Handbook of perception and human performance (Vol. 1). New York: Wiley.

Zahabi, S. J. N., Islam, M. S., Kim, S., Lau, N., Nussbaum, M. A., & Lim, S. (2025). Cognitive workload assessment during VR forklift training. International Journal of Industrial Ergonomics, 107, 103718.

Naseri, N., Parastesh, F., Ghassemi, F., Jafari, S., Perc, M., & Završnik, J. (2025). Synchronization levels in EEG connectivity during cognitive workloads while driving. Nonlinear Dynamics, 113(7), 7243-7258.

Qin, Y., & Bulbul, T. (2023). An EEG-based mental workload evaluation for AR head-mounted display use in construction assembly tasks. Journal of Construction Engineering and Management, 149(9), 04023088.

Adachi, M., Nobukawa, S., & Inagaki, K. (2024). Assessment of driver fatigue-related brain responses and causal factors during driving under different traffic conditions. Frontiers in Applied Mathematics and Statistics, 10, 1426253.

Ji, L., Zhang, C., Li, H., Zhang, N., Guo, C., Zhang, Y., & Tang, X. (2023). A comprehensive experimental framework based on analysis of the pilot's EEG and NASA-TLX questionnaire in a VR environment. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, 237(7), 869-878.

Hutson, J. W., Franklin, A. E., Rogers, B. A., & Walker, D. (2024). Psychometric Testing of NASA-TLX to Measure Learners' Cognitive Load in Individual and Group Nursing Simulations. Clinical Simulation in Nursing, 95, 101607.

Jajo, N., & Biondi, F. N. (2023, September). Using detection response task and NASA-TLX to measure the difference in cognitive workload between partially automated mode and manual mode: an on-road study. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 67, No. 1, pp. 2508-2512). Sage CA: Los Angeles, CA: SAGE Publications.

Arana-De las Casas, N. I., De la Riva-Rodríguez, J., Maldonado-Macías, A. A., & Sáenz-Zamarrón, D. (2023). Cognitive analyses for interface design using dual n-back tasks for mental workload (MWL) evaluation. International Journal of Environmental Research and Public Health, 20(2), 1184.

Yeh, P. C., & Chao, C. J. (2024). Effects of simulated driving task screen background color combinations on visual performance and visual fatigue. Journal of the Chinese Institute of Engineers, 47(8), 970-976.

Javernik, A., Buchmeister, B., & Ojsteršek, R. (2023). The NASA-TLX approach to understanding workers' workload in human-robot collaboration. International journal of simulation modelling, 22(4), 574-585.

Özdemir, M., Dönmez, K., & Demirel, S. (2023). Determining the key factors affecting pseudo-pilot workload based on real-time simulations. Transportation Research Record, 2677(7), 718-732.

Ruan S, Lin B, Zhu Y, *et al*. (2022). EEG power spectrum of different mental loads in simulated flight tasks[J]. Military medicine, 46(07):499-504.

Borghini, G., Vecchiato, G., Toppi, J., Astolfi, L., Maglione, A., Isabella, R., ... & Babiloni, F. (2012, August). Assessment of mental fatigue during car driving by using high-resolution EEG activity and neurophysiologic indices. In the 2012 annual international conference of the IEEE engineering in medicine and biology society (pp. 6442-6445). IEEE.

Hebbar, P. A., Bhattacharya, K., Prabhakar, G., Pashilkar, A. A., & Biswas, P. (2021). Correlation between physiological and performance-based metrics to estimate pilots' cognitive workload. Frontiers in psychology, 12, 555446.

Pavlov, Y. G., & Kotchoubey, B. (2017). EEG correlates of working memory performance in females. BMC Neuroscience, 18, 1-14.

Biernacki, M. P., & Lewkowicz, R. (2024). The role of visual conditions and aircraft type on different aspects of pilot workload. Applied Ergonomics, 118, 104268.

Vidulich, M. A., & Tsang, P. S. (2012). Mental workload and situation awareness. Handbook of human factors and ergonomics, 243-273.

Holm, A., Lukander, K., Korpela, J., Sallinen, M., & Müller, K. M. (2009). Estimating brain load from the EEG. The Scientific World Journal, 9(1), 639-651.

Li, W., Cheng, S., Dai, J., & Chang, Y. (2025). Effects of Mental Workload Manipulation on Electroencephalography

Spectrum Oscillation and Microstates in Multitasking Environments. Brain and Behavior, 15(1), e70216.

Yu, H., Cao, W., Fang, T., Jin, J., & Pei, G. (2024). EEG β oscillations in aberrant data perception under cognitive load modulation. Scientific Reports, 14(1), 22995.

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.

He, D., Donmez, B., Liu, C. C., & Plataniotis, K. N. (2019). High cognitive load assessment in drivers through wireless electroencephalography and the validation of a modified N-back task. IEEE Transactions on Human-Machine Systems, 49(4), 362-371.

Örün, Ö., & Akbulut, Y. (2019). Effect of multitasking, physical environment, and electroencephalography use on cognitive load and retention. Computers in Human Behavior, 92, 216-229.

Wilson, G. F. (2002). An analysis of mental workload in pilots during flight using multiple psychophysiological measures. The International Journal of Aviation Psychology, 12(1), 3-18.

Rozado, D., & Dunser, A. (2015). Combining EEG with pupillometry to improve cognitive workload detection. Computer, 48(10), 18-25.

Hu, Y., Ni, Q., & Lü, W. (2022). Avoidant personality disorder symptoms and cardiovascular reactivity to psychological stress tasks with increasing cognitive demands. Journal of Psychophysiology.

Chanel, G., Kronegg, J., Grandjean, D., & Pun, T. (2006, September). Emotion assessment: Arousal evaluation using EEGs and peripheral physiological signals. In International workshop on multimedia content representation, classification and security (pp. 530-537). Berlin, Heidelberg: Springer Berlin Heidelberg.

Wilson, G. F., & Russell, C. A. (2007). Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. Human factors, 49(6), 1005-1018.

Nilsson, E. J., Bärgman, J., Ljung Aust, M., Matthews, G., & Svanberg, B. (2022). Let complexity bring clarity: A multidimensional assessment of cognitive load using physiological measures. Frontiers in neuroergonomics, 3, 787295.

Hogervorst, M. A., Brouwer, A. M., & Van Erp, J. B. (2014). Combining and comparing EEG, peripheral physiology, and eye-related measures for the assessment of mental workload. Frontiers in neuroscience, 8, 322.

Brouwer, A. M., Hogervorst, M. A., Oudejans, B., Ries, A. J., & Touryan, J. (2017). EEG and eye tracking signatures of target encoding during structured visual search. Frontiers in human neuroscience, 11, 264.

Jiang, Z., Li, X., Ge, L., Xu, J., Lu, Y., Zhang, Y., & Mao, M. (2024). Using multimodal methods and machine learning to recognize mental workload: Distinguishing between underload, moderate load, and overload. International Journal of Human–Computer Interaction, 1-17.

Parmentier, D. D., Van Acker, B. B., Detand, J., & Saldien, J. (2020). Design for assembly meaning: a framework for designers to design products that support operator cognition during the assembly process. Cognition, Technology & Work, 22, 615-632.

Jin, H., Liu, L., Luo, Z., Meng, S., & Zhao, Y. (2024). The effects of different interruption conditions on mental workload: an experimental study based on multimodal measurements. Ergonomics, 1-19.

Chen, F., Zhou, J., Wang, Y., Yu, K., Arshad, S. Z., Khawaji, A., ... & Conway, D. (2016). Theoretical aspects of multimodal cognitive load measures. Robust Multimodal Cognitive Load Measurement, 33-71.

Longo, F., Nicoletti, L., & Padovano, A. (2017). Smart operators in Industry 4.0: A human-centered approach to enhance operators' capabilities and competencies within the new smart factory context. Computers & Industrial Engineering, 113, 144-159.

Landi, C. T., Villani, V., Ferraguti, F., Sabattini, L., Secchi, C., & Fantuzzi, C. (2018). Relieving operators' workload: Towards affective robotics in industrial scenarios. Mechatronics, 54, 144-154.

Chan, W. P., Hanks, G., Sakr, M., Zhang, H., Zuo, T., Van der Loos, H. M., & Croft, E. (2022). Design and evaluation of an augmented reality head-mounted display interface for human-robot teams collaborating in physically shared manufacturing tasks. ACM Transactions on Human-Robot Interaction (THRI), 11(3), 1-19.

Demerouti, E. (2022). Turn digitalization and automation to a job resource. Applied Psychology, 71(4), 1205-1209.