

PILOT WORKLOAD PROFILES UNDER CONTINUOUS MULTI-TASKING AND SPARSE EVENTS VIGILANCE

Wenduo Xiao[1], Chunling Zhao[2], Yang Gao[1], Yanyu Lu[1] & Shan Fu[1]

[1]School of Automation and Intelligent Sensing, Shanghai Jiao Tong University, Shanghai, China
[2]Shanghai Aircraft Design and Research Institute, Shanghai, China

Abstract

As aircraft automation advances, pilot activity is shifting from continuous manual control toward supervisory monitoring and response to sparse, time-critical events. This study compares workload characteristics under two task structures: continuous multi-task control, implemented with a Multi-Attribute Task Battery (MATB-II) paradigm, and sparse event vigilance, implemented with a flight-simulator-based radar vigilance paradigm. Ten adult participants completed both paradigms with concurrent acquisition of behavioral, ocular, and electroencephalographic measures. Evidence from both paradigms suggests that workload profiles are shaped by task temporal structure. Continuous multi-task control was associated with resource-limited execution, selective performance degradation, attentional narrowing toward high-priority components, and secondary frontal spectral shifts. Sparse event vigilance preserved or improved behavioral response efficiency in some conditions, yet was accompanied by ocular modulation and descriptive Hjorth-based neurophysiological patterns. These findings suggest that workload in highly automated flight decks should not be treated as a single scalar quantity, and that future human-machine interfaces should account for covert cognitive costs during low-activity monitoring periods.

Keywords: pilot workload, continuous multi-task control, sparse event vigilance, EEG, eye tracking

1. Introduction

Increasing automation in modern aircraft is reshaping the temporal structure of pilot work. In conventional manual or semi-automated operations, pilots remain continuously engaged in control, monitoring, communication, and system management. In highly automated contexts, the operator role shifts toward supervisory monitoring, diagnosis, and intervention upon discrete abnormal or safety-critical events [1, 2]. This transition produces a workload profile that is poorly captured by methods developed for sustained, continuously active task performance.

This distinction matters because workload is not merely a matter of task quantity or event rate. Continuous multi-task control and sparse event vigilance impose qualitatively different demands on attention, readiness, and response preparation. Continuous multi-task control requires the ongoing allocation of limited cognitive resources across concurrent subtasks. Sparse event vigilance, in contrast, may entail prolonged periods of low overt activity while still demanding sustained readiness for rare but consequential events [3, 4]. In highly automated flight decks, this can produce a dissociation between observable performance and the underlying workload state.

Prior aviation workload research has extensively examined continuous task performance and continuous multi-task control through behavioral performance, subjective ratings, ocular measures, and neurophysiological indices [5, 6]. However, few studies have directly contrasted continuous multi-task control with sparse event vigilance within a unified assessment framework. This omission constrains the interpretation of workload metrics in automation-rich operations, where adequate performance during sparse monitoring may mask elevated cognitive effort, anticipatory control, or stress.

The present study addresses this gap by comparing workload profiles under two task structures: a continuous multi-task control condition based on the Multi-Attribute Task Battery (MATB-II), and a sparse event vigilance condition designed to approximate supervisory monitoring in highly automated flight operations. Behavioral, ocular, and electroencephalographic (EEG) measures were collected to characterize both overt performance and covert physiological modulation. The central question is whether sparse event vigilance constitutes a simple reduction in workload relative to continuous multi-task control, or whether it represents a qualitatively distinct workload state.

The contributions of this paper are threefold. First, it contrasts continuous multi-task control and sparse event vigilance using a common multimodal assessment framework. Second, it identifies behavioral and physiological signatures that diverge across task structures. Third, it discusses implications for workload-adaptive human-machine interface (HMI) design in future highly automated flight decks.

2. Related Work

2.1 Pilot Workload in Automated Flight Operations

Automation reduces manual control demand while increasing the importance of monitoring, mode awareness, and timely intervention. This shift has been linked to automation surprises, complacency, and monitoring failures when operators must supervise reliable systems over extended periods [1, 2]. Under such conditions, low overt activity should not be equated with the absence of cognitive demand. Rather, the operator must maintain readiness, detect weak cues, and prepare for infrequent events of high consequence.

Recent reviews of pilot workload assessment emphasize the necessity of multimodal measurement, since no single indicator fully captures cognitive state across operational contexts [5, 6]. Behavioral performance is essential but may be insensitive to compensatory effort, particularly when participants preserve accuracy or response speed by recruiting additional cognitive resources. Physiological and ocular measures therefore provide complementary information about hidden workload costs.

2.2 Continuous Multi-Task Control

The MATB-II is widely used to elicit and study multi-task workload in aviation-relevant experimental settings. It requires participants to divide attention across subtasks including tracking, system monitoring, resource management, and communications [7]. Because the subtasks are concurrent and demand repeated action, MATB-II performance is well suited for investigating resource allocation, attentional tunneling, and degradation under increasing continuous demand.

In continuous multi-task control, increased workload typically produces observable performance costs, but these costs are often unevenly distributed across subtasks. Some subtasks remain stable due to assigned priority or continuous visibility, whereas peripheral or intermittent monitoring subtasks may suffer omission errors. This pattern is directly relevant to flight deck design, since pilots may preserve primary control at the expense of less salient yet operationally critical information.

2.3 Sparse Event Vigilance

Sparse event vigilance tasks demand sustained attention to detect infrequent and sometimes unpredictable signals. Classical vigilance research demonstrates that this form of work can be mentally demanding and stressful even when external activity is minimal [3, 8]. In aviation automation, sparse events may include abnormal alerts, traffic conflicts, system mode changes, or other discrete conditions requiring rapid intervention.

Sparse event vigilance differs from continuous multi-task control in that task demand is structured around readiness rather than uninterrupted execution. Higher event frequency may enhance alertness or response preparation under some conditions, whereas very low frequency may foster disengagement or missed detections. Workload in sparse event vigilance therefore cannot be inferred from the number of actions performed alone; it must be evaluated jointly through performance and physiological state.

2.4 Physiological and Ocular Workload Measures

EEG spectral measures have been widely used to evaluate cognitive workload, with particular emphasis on frontal low-frequency and beta-band activity [9]. Relative power across delta, theta, beta, and gamma bands provides complementary information about how the spectral composition of EEG changes across task states. Temporal descriptors such as Hjorth Mobility and Complexity complement spectral analysis by capturing signal variance and changes in frequency structure over time [10]. Eye-tracking measures, including blink rate, saccade rate, fixation allocation, and pupil diameter, provide additional information on arousal and visual attention distribution.

For the present comparison, these measures are valuable because continuous task engagement and sparse readiness may yield distinct physiological signatures even when behavioral performance does not degrade in the same manner. A multimodal approach is thus required to disentangle workload magnitude from workload type.

3. Experimental Design

3.1 Participants and Ethics

Ten adult volunteers participated in the study, with a mean age of 24 years. All participants had completed several rounds of prior task training in related experiments, and all enrolled participants were retained for analysis. Per institutional requirements, ethical approval was not required for this minimal-risk study, which used non-invasive EEG and eye-tracking measurements during routine experimental tasks. Written informed consent was obtained prior to participation, and the collected data were anonymized and used exclusively for research purposes.

3.2 Overall Design

A within-subject design was adopted to compare neurophysiological, ocular, and behavioral responses under two task structures: continuous multi-task control and sparse event vigilance. Each participant completed both paradigms. The continuous multi-task control condition represented sustained active engagement, whereas the sparse event vigilance condition represented supervisory monitoring with intermittent event-response requirements.

The two paradigms were chosen to instantiate two operationally important temporal structures of pilot work via distinct task surfaces. The MATB-II paradigm isolates the effect of serial versus parallel execution of an identical subtask set. The sparse event vigilance paradigm embeds intermittent target detection within a flight-simulator task, thereby approximating the monitoring and readiness burden of automation-rich operations.

The eye-tracking analyses retained complete paired cases for each contrast, yielding 10 participants for the reported inferential tests. EEG spectral results are reported as condition-level relative band-power summaries across the vigilance and MATB-II states.

3.3 Continuous Multi-Task Control Paradigm

The continuous multi-task control condition was implemented using MATB-II. Participants performed four aviation-relevant subtasks: compensatory tracking, system monitoring, resource management, and communications. The tracking task required participants to keep a randomly moving target near the display center using a joystick or mouse. The system monitoring task required detection of abnormal indicator or scale states. The resource management task required operation of pump switches to maintain virtual tank levels. The communications task required participants to monitor auditory messages, identify calls directed to themselves, and tune the correct radio frequency. The MATB-II interface is shown in Figure 1.

The low-load MATB-II condition employed serial execution: participants completed the four subtasks independently and sequentially, thereby reducing between-task interference and providing a baseline of subtask-specific workload. The high-load MATB-II condition employed parallel execution, with participants handling the same four subtasks simultaneously, thereby increasing demand on divided attention, task switching, parallel processing, and working-memory updating. Since the subtask set was held constant, the principal experimental contrast was the additional cognitive cost of concurrent execution.

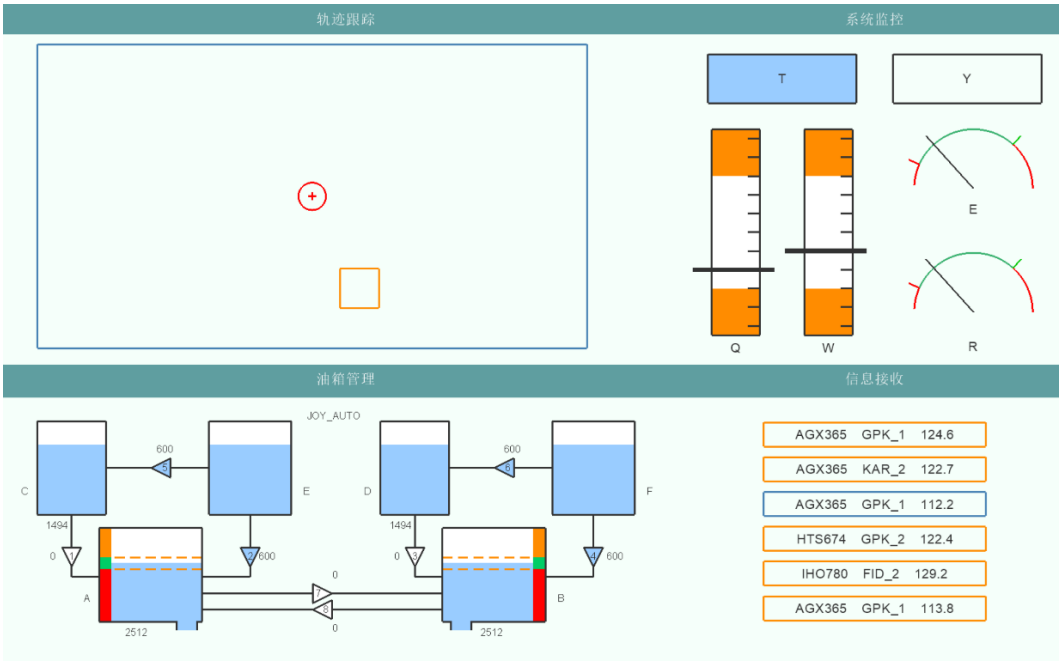


Figure 1 – MATB-II interface used for the continuous multi-task control paradigm.

The MATB-II procedure was standardized. Participants first read the task instructions and provided informed consent. They then practiced each of the four subtasks until they understood the task requirements and reached the predetermined familiarity criterion. Physiological recording devices were subsequently fitted, and signal quality was verified. Prior to formal task execution, a five-minute eyes-open resting baseline was recorded while participants fixated on a central cross. Participants then completed the low- and high-load MATB-II task blocks, with block order counterbalanced across participants where feasible and a 5–10 minute rest interval between blocks.

The analysis is based on platform-derived composite and subtask scores for the serial and parallel MATB-II blocks. The serial and parallel task labels define the implemented workload contrast, and the automatic MATB-II summary scores serve as performance indicators for that contrast.

3.4 Sparse Event Vigilance Paradigm

The sparse event vigilance condition was conducted in a fixed-base flight simulator at the Human-Machine and Environmental Intelligence Systems Laboratory of Shanghai Jiao Tong University. The simulator provided a cockpit layout with flight stick, throttle, pedals, instrument displays, and a wide-angle external visual system. The platform was integrated with multimodal recording equipment, including the EEG and eye-tracking devices used in the present analysis; heart-rate and fNIRS signals fell outside the analytic scope of this paper. The simulator environment and flight display are shown in Figure 2.

The vigilance task combined a primary flight task with a secondary radar monitoring task. In the primary task, participants used the flight controls to maintain attitude, altitude, and route progress. The flight scenario required travel from the initial position toward the SUF area, followed by a turn toward the VMB area. The flight was to be completed within five minutes while maintaining altitude between 45,000 ft and 47,000 ft. Participants monitored the primary flight display and navigation display to maintain altitude, vertical speed, heading, and route alignment.



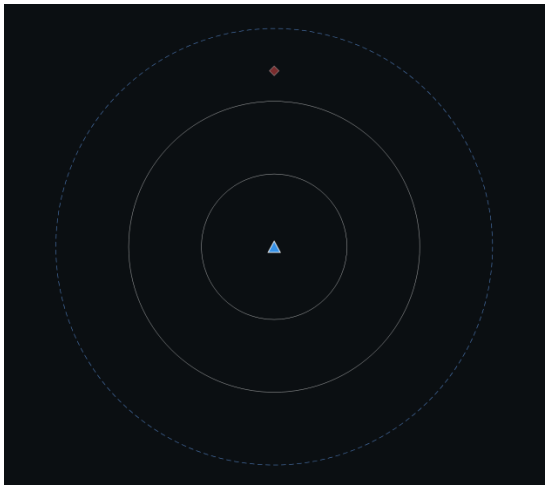
(a) Simulator setup.



(b) Flight display.

Figure 2 – Experimental environment for the sparse event vigilance paradigm.

In the secondary radar task, participants monitored a situation display while continuing the flight task. The ownship was represented by a blue triangle at the center of the radar interface. Enemy aircraft appeared from the outer visual reference circle and moved toward or across the ownship area. Participants were instructed to press the “L” key when an enemy aircraft entered the designated alert reference circle. The system recorded response time, response correctness, and the target position at the moment of response. Figure 3 illustrates the radar display and target appearances.



(a) Radar interface.



(b) Type A.



(c) Type B.



(d) Type C.



(e) Type D.

Figure 3 – Radar display and target appearances for the sparse event vigilance task.

The radar task comprised both static and dynamic elements. Static elements included the ownship symbol, the outer visual reference circle, and concentric alert reference circles. Dynamic elements consisted of diamond-shaped enemy aircraft. Enemy motion followed one of three patterns: straight crossing at constant speed, arc-like fly-by motion, or intercept-like motion with continuous turning toward the ownship until passing a distance threshold. The visual form of enemies varied in color and size, yielding four target appearances used in the radar interface.

Difficulty in the sparse event vigilance paradigm was manipulated through task environment and event frequency. The flight-only baseline condition consisted of flight simulation only: participants

performed the primary flight task, but no radar alert targets were presented and no radar-response keypress was required. This baseline condition provides the reference flight state included in the subsequent physiological figures. In the low-load condition, participants flew under calm conditions without crosswind while responding to low-frequency enemy appearances (approximately 2–3 events per minute). In the high-load condition, participants maintained route stability under random crosswind disturbance while responding to higher-frequency enemy appearances (approximately 5–6 events per minute). The task design also incorporated a comparison in which enemy velocity was doubled while other conditions were held constant; velocity is treated here as a secondary manipulation in the present condition definitions. Table 1 summarizes the resulting condition definitions.

The complete vigilance session comprised an initial flight-only baseline run, followed by two flight-plus-vigilance runs corresponding to the low- and high-load conditions. Participants were trained on the flight controls and radar task prior to formal data collection. They were then fitted with physiological recording devices, completed the baseline flight, rested briefly, and completed the two vigilance flights with rest between runs.

The available task records support describing target velocity and motion pattern as design elements. The results therefore focus on the low- versus high-difficulty vigilance contrast, with target velocity and motion pattern treated as secondary task characteristics rather than inferential factors.

Table 1 – Task condition definitions used in the two paradigms.

Paradigm	Condition	Task structure	Workload interpretation
MATB-II	Low load	Four subtasks executed serially and independently.	Baseline load for each subtask with limited resource competition.
MATB-II	High load	The same four subtasks executed simultaneously.	Continuous divided-attention demand and resource competition.
Sparse event vigilance	Flight-only baseline	Flight simulation only, with no radar alert targets and no radar-response keypress requirement.	Reference flight state without radar monitoring demand.
Sparse event vigilance	Low load	Calm flight plus low-frequency radar events.	Supervisory monitoring with lower event density.
Sparse event vigilance	High load	Random crosswind plus higher-frequency radar events.	Combined flight-control disturbance and increased event response demand.

4. Data Acquisition and Metrics

4.1 Behavioral Measures

Behavioral performance metrics included response time, accuracy, hit rate, omission errors, and commission errors. In the MATB-II condition, behavioral measures were analyzed at the subtask level to determine whether increased continuous load differentially affected tracking, monitoring, resource management, and communications. The MATB-II platform automatically logged operation behavior and performance data, including tracking error, system monitoring response time and accuracy, resource management performance, and communication accuracy.

For the sparse event vigilance condition, behavioral measures were analyzed as a function of event frequency and target velocity. Response time was defined as the interval between an enemy aircraft entering the alert reference circle and the participant's keypress. A hit was defined as a valid response after alert entry within the accepted response window. An omission error was defined as a failure to respond to a target requiring response. A commission error was defined as a response issued in the absence of a valid target.

The behavioral vigilance figures characterize the observed response-time and reaction-type patterns across task settings. Paired inferential tests are reported for the eye-tracking measures, while EEG spectral measures are used as condition-level relative-power descriptors.

4.2 Eye-Tracking Measures

Eye-tracking measures characterized arousal and visual attention allocation. The principal indicators were blink rate, saccade rate, pupil diameter, fixation distribution, and display-area sampling. For the MATB-II condition, gaze observations were used to characterize attentional narrowing toward central or high-priority task components under concurrent demand. For the sparse event vigilance condition, ocular measures were used to examine monitoring strategy across the primary flight display, navigation display, and radar interface.

The eye tracker employed in this study was the aSee Glasses. Raw gaze and pupil streams were summarized as valid-gaze fraction, mean and median pupil diameter, eyelid openness, blink rate, blink fraction, normalized gaze speed, and a saccade-rate proxy. Missing gaze was reflected in the valid-gaze fraction, and per-sample summary metrics were computed from available valid samples. Whole-stream summaries were used for the paired statistical tables, while AOI fixation observations supported interpretation of attention allocation in the continuous multi-task control task.

4.3 EEG Measures

EEG recordings were analyzed using both temporal and spectral approaches. Temporal dynamics were quantified using Hjorth Activity, Mobility, and Complexity. Activity reflects signal variance, Mobility reflects the normalized rate of signal change, and Complexity reflects changes in frequency structure. For a signal $x(t)$ with first derivative $\dot{x}(t)$, Activity is defined as

$$Activity = var(x(t)). \tag{1}$$

Mobility and Complexity are defined as

$$Mobility = \sqrt{\frac{var(\dot{x}(t))}{var(x(t))}}, \quad Complexity = \frac{Mobility(\dot{x}(t))}{Mobility(x(t))}. \tag{2}$$

Frequency-domain analyses focused on frontal relative delta, theta, beta, and gamma band activity to support interpretation of the temporal features. These EEG measures were selected because they capture changes in neural engagement that may not manifest directly in behavioral performance.

Spectral EEG measures were derived from BDF recordings sampled at 1000 Hz. Relative band powers were computed from Welch spectra and averaged over frontal channels Fp1, Fp2, Fz, F3, F4, F7, F8, FC1, FC2, FC5, and FC6 when present. The analyzed bands were delta, theta, beta, and gamma. The spectral results are interpreted as condition-level frontal relative band-power summaries rather than baseline-normalized event-related EEG effects.

4.4 Statistical Analysis

Condition effects with complete paired data were evaluated using paired-sample t tests. The two reported contrasts were continuous multi-task control serial versus parallel execution and sparse event vigilance low versus high difficulty. The eye-tracking tests retained 10 participants. Mean differences and percentage changes are reported as descriptive effect-size summaries, and exact p values are interpreted with caution given the small sample and the absence of formal multiple-comparison correction.

For the MATB-II descriptive table, the percentage decline from serial to parallel execution was computed as

$$D_m = \frac{S_m - P_m}{S_m} \times 100\%, \tag{3}$$

where S_m and P_m denote the serial and parallel mean values for metric m, respectively. This descriptive calculation is used solely to summarize the MATB-II performance table and should not be interpreted as an inferential test.

The results separate the paired statistical tests reported for eye-tracking metrics from the descriptive EEG spectral and Hjorth patterns shown in the task figures. This separation keeps the quantitative paired tests and the figure-based workload interpretation explicit.

5. Results

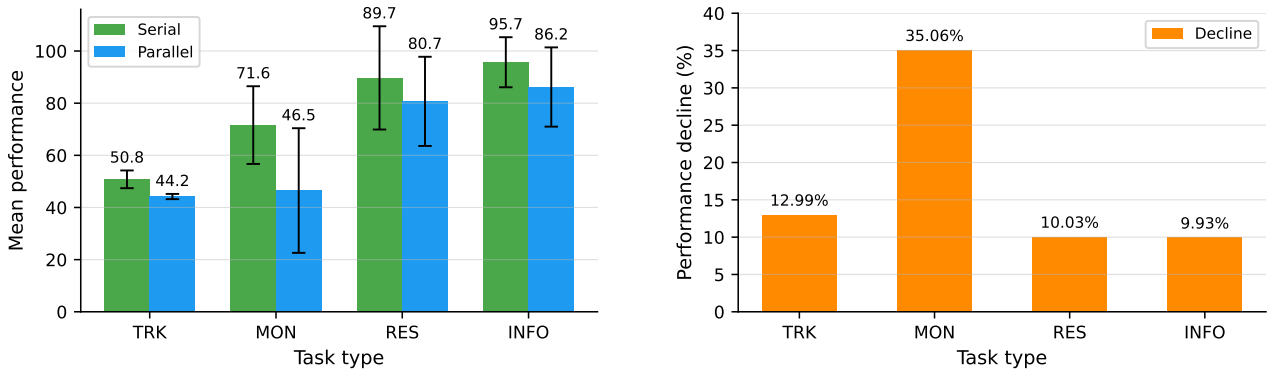
5.1 Continuous Multi-Task Control

The MATB-II results show the expected cost of concurrent task execution. As summarized in Table 2 and Figure 4, the largest subtask decline occurred in system monitoring, falling from 71.6 to 46.5 with a 35.06% decline indicator. Tracking decreased from 50.8 to 44.2, while resource management and communications remained comparatively high, declining from 89.7 to 80.7 and from 95.7 to 86.2, respectively.

Table 2 – MATB-II performance under serial and parallel task execution.

Metric	Serial mean	Parallel mean	Decline
Tracking	50.8	44.2	12.99%
System monitoring	71.6	46.5	35.06%
Resource management	89.7	80.7	10.03%
Communications	95.7	86.2	9.93%

The decline pattern is visualized in Figure 4. Crucially, continuous multi-task control yielded selective rather than uniform degradation across subtasks. Participants preserved some subtasks better than others, while performance on monitoring-related components declined more steeply. This pattern is consistent with resource-limited execution under concurrent demand: participants protect central or immediately actionable components, leaving peripheral or intermittent monitoring more vulnerable. The selective task-performance decline is consistent with attentional narrowing observed in the AOI fixation pattern from the eye-tracking records.



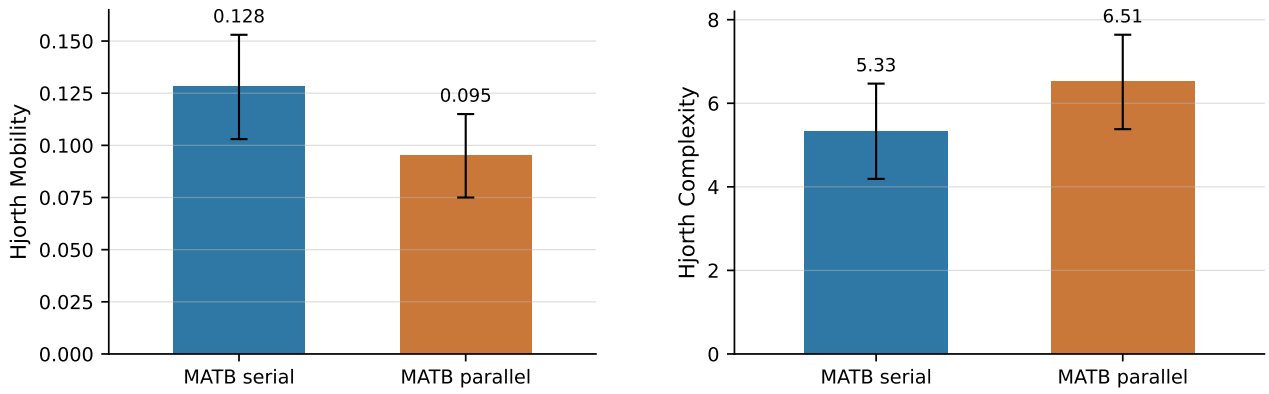
(a) Task performance.

(b) Decline under continuous multi-task control.

Figure 4 – Behavioral performance under serial and parallel MATB-II conditions.

Aggregate eye-tracking summaries also shifted between the serial and parallel continuous multi-task control blocks. From serial to parallel execution, valid-gaze fraction decreased by 1.2%, mean pupil diameter increased by 0.6%, blink fraction increased by 15.5%, mean gaze speed decreased by 4.3%, the 95th-percentile gaze speed decreased by 16.9%, and the saccade-rate proxy decreased by 12.6%. The saccade-rate proxy showed the clearest paired shift among these whole-stream eye metrics. These shifts provide ocular-state context for the attentional narrowing suggested by the selective MATB-II performance decline and AOI fixation observations.

The EEG Hjorth results provide descriptive physiological context for serial and parallel MATB-II execution. Figure 5 shows condition-level differences in Hjorth Mobility and Complexity. These observations are compatible with the interpretation that continuous multi-task control may alter neural signal dynamics alongside overt task performance, but they should be treated as supportive rather than primary evidence.



(a) Hjorth Mobility.

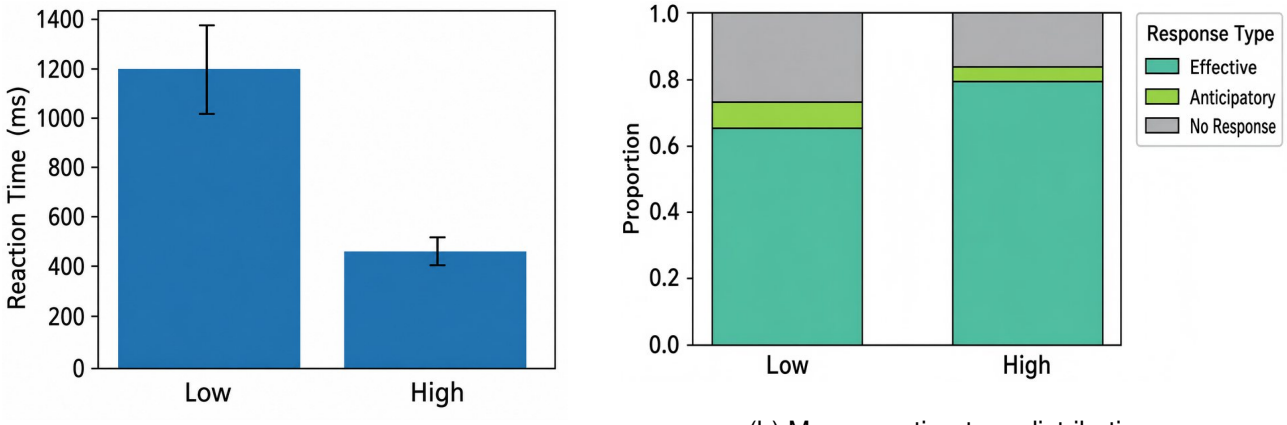
(b) Hjorth Complexity.

Figure 5 – EEG Hjorth features under serial and parallel MATB-II conditions.

Figure 8 provides the frontal relative-power context for the continuous multi-task control contrast. Compared with serial MATB-II execution, parallel execution preserved theta at a similar level (0.076 versus 0.075), showed higher delta (0.354 to 0.364) and beta (0.111 to 0.125), and showed lower gamma (0.172 to 0.137). This pattern suggests a modest shift in frontal spectral composition under parallel MATB-II execution. Together with the Hjorth features, these spectral results are interpreted as secondary descriptive context for the continuous multi-task control contrast.

5.2 Sparse Event Vigilance

In the sparse event vigilance paradigm, higher task difficulty was operationalized through increased event frequency and, where applicable, faster target motion. Figure 6 shows a performance–load separation in the behavioral response patterns. Mean response time varied systematically across task settings, and the distribution of reaction types shifted with difficulty. Higher event density may enhance readiness and response preparation, yielding faster or more efficient responses in some conditions even when the monitoring state remains physiologically demanding.



(a) Mean response time.

(b) Mean reaction-type distribution.

Figure 6 – Behavioral response patterns in the sparse event vigilance task.

The paired eye-tracking results in Figure 7 show clearer modulation for the low- versus high-difficulty sparse event vigilance contrast than for the serial- versus parallel-execution continuous multi-task control contrast. From low to high difficulty, valid-gaze fraction decreased from 0.913 to 0.894. Blink rate increased from 24.58 to 27.69 blinks/min. Mean normalized gaze speed increased by 21.4%; the 95th-percentile gaze speed increased by 47.2%; and the saccade-rate proxy increased from 184.20 to 204.00 events/min. Mean pupil diameter remained essentially stable, from 5.397 to 5.394 mm. This pattern aligns more closely with altered visual scanning and arousal than with a pupil-diameter workload effect.

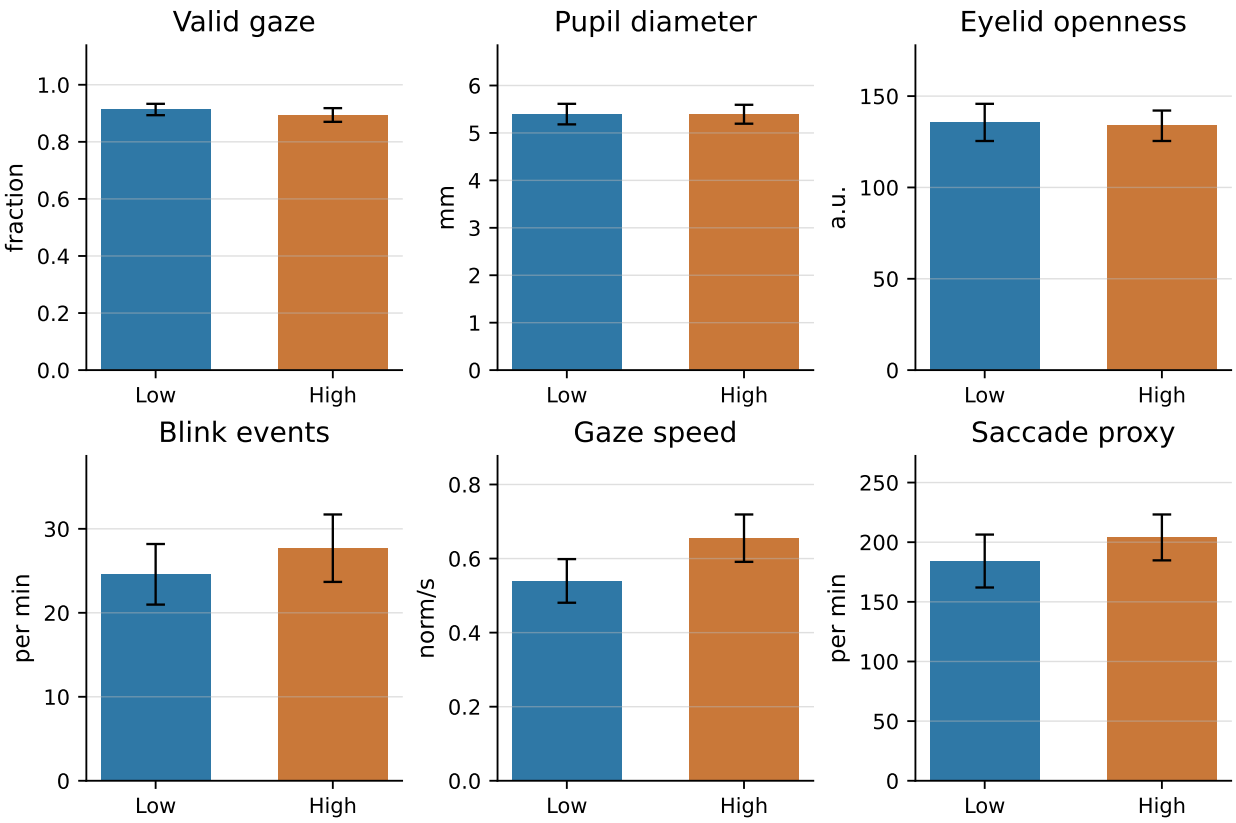


Figure 7 – Eye-tracking metrics under low- and high-difficulty sparse event vigilance.

Figure 8 also provides frontal spectral context for sparse event vigilance. Relative gamma was higher in the low- and high-difficulty vigilance runs than in the vigilance baseline (0.296 and 0.281 versus 0.266), whereas delta remained close across the low- and high-difficulty runs (0.304 and 0.303). From low to high difficulty, theta increased slightly from 0.061 to 0.064 and beta increased from 0.089 to 0.097, while gamma decreased from 0.296 to 0.281. The spectral pattern is interpreted as secondary context complementing the ocular shifts and descriptive Hjorth trends: sparse event vigilance runs showed elevated high-frequency relative power compared with baseline, while the high-difficulty run introduced modest theta and beta increases relative to the low-difficulty run.

The EEG Hjorth features in Figure 9 provide descriptive temporal-dynamics context for the sparse event vigilance conditions. The figure shows opposing trends for Mobility and Complexity across difficulty settings, consistent with the possibility that sparse event vigilance changes neural dynamics rather than merely scaling a single workload level. Increased event density was accompanied by elevated Hjorth Complexity and a slight reduction in Hjorth Mobility, a pattern compatible with sustained readiness and anticipatory processing during monitoring.

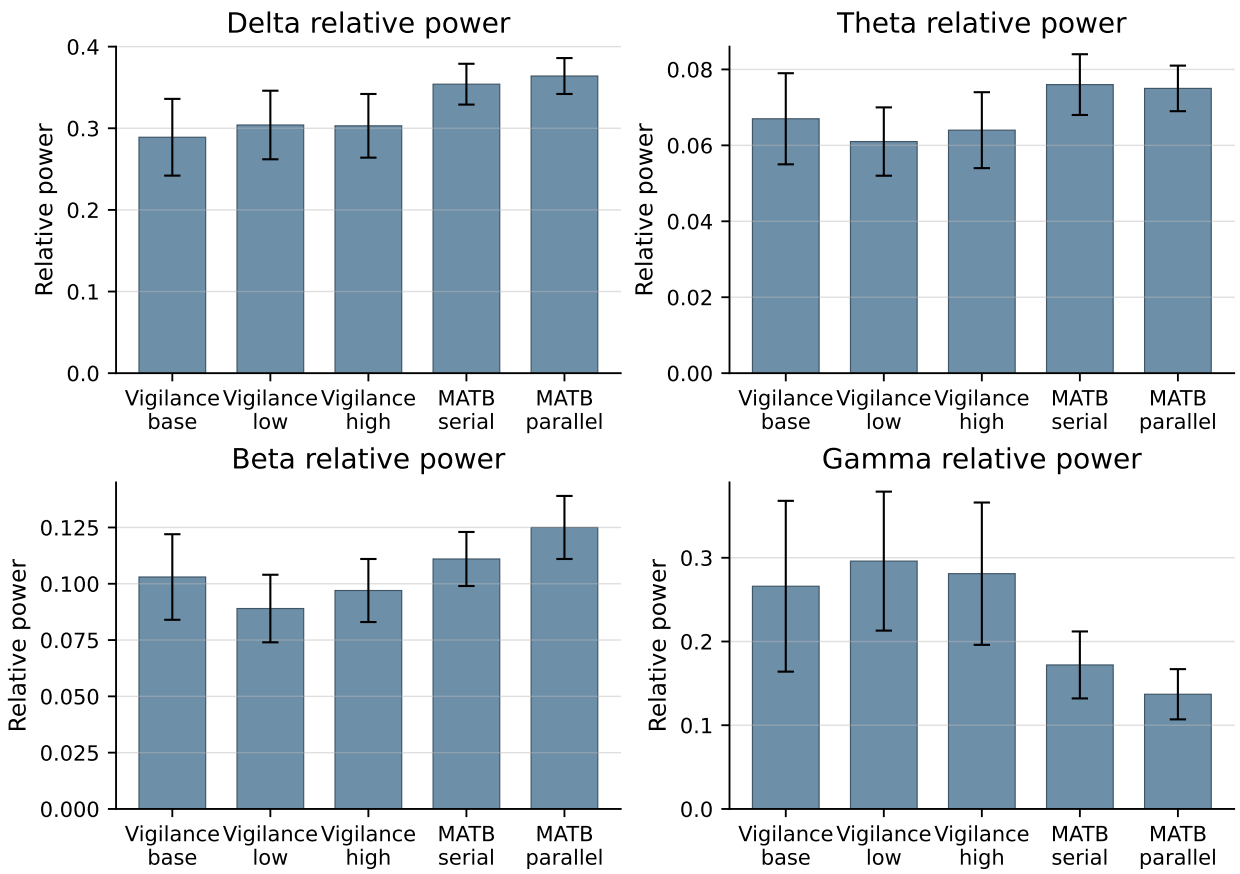
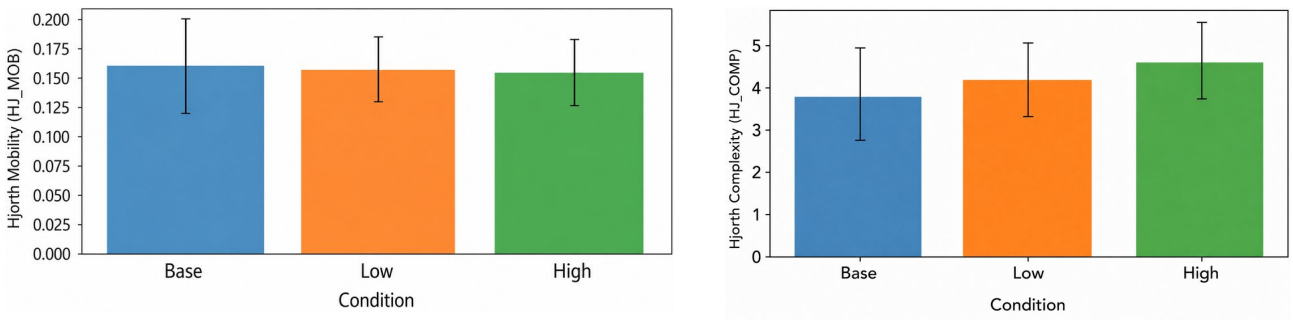


Figure 8 – Frontal EEG relative band-power comparisons across vigilance and MATB-II conditions.



(a) Hjorth Mobility.

(b) Hjorth Complexity.

Figure 9 – EEG Hjorth features under sparse event vigilance difficulty settings.

Taken together, Figures 6–9 suggest a dissociation between behavioral performance and physiological activity in the sparse event vigilance task. Behavioral metrics indicate maintained or improved task efficiency under higher event density, whereas eye-tracking changes and Hjorth feature patterns are consistent with an altered monitoring state characterized by sustained readiness and anticipatory control.

5.3 Cross-Paradigm Comparison

Direct comparison of the two paradigms suggests that workload profiles are shaped by task temporal structure. Continuous multi-task control primarily reflected resource-limited execution under concurrent demand, producing selective performance degradation and attentional narrowing toward high-priority task components. Sparse event vigilance appeared to involve a qualitatively different form of workload, characterized by sustained readiness and anticipatory control that could coexist with maintained or slightly improved behavioral performance.

This contrast is important because nominally similar workload labels may denote distinct operational states. In the continuous multi-task control condition, workload manifested through competition among simultaneous actions and selective performance loss. In the sparse event vigilance condition, preserved performance together with ocular modulation and descriptive Hjorth trends is more consistent with anticipatory readiness and arousal than with the resource-depleting competition seen in multitasking. Table 3 summarizes the paired eye-tracking comparisons together with descriptive EEG spectral and Hjorth summaries, and Table 4 summarizes the resulting workload signatures. A workload assessment system that treats these states as a single continuum risks confounding overt task failure with covert readiness cost.

Table 3 – Physiological summaries for continuous multi-task control and sparse event vigilance.

Task contrast	Key observations	Interpretation
Continuous multi-task control: serial vs parallel execution	Eye metrics shifted modestly: valid gaze -1.2%, pupil diameter $+0.6\%$, blink fraction $+15.5\%$, and saccade proxy -12.6%. Frontal delta and beta increased slightly, gamma decreased, and Hjorth Mobility/Complexity showed descriptive directional changes.	Behavioral performance and AOI observations carry the main continuous-load interpretation; EEG spectral and Hjorth results are secondary context.
Sparse event vigilance: low vs high difficulty	Ocular shifts were clearer: valid gaze -2.1%, blink rate $+12.6\%$, gaze speed $+21.4\%$, and saccade proxy $+10.8\%$. Theta and beta increased slightly; Hjorth changes were descriptive.	Ocular metrics provide the strongest support for altered scanning and arousal; EEG spectral and Hjorth trends remain supportive context.

Table 4 – Summary of workload signatures across the two task structures.

Dimension	Continuous multi-task control	Sparse event vigilance
Temporal structure	Repeated concurrent actions across subtasks.	Long monitoring periods interrupted by discrete events.
Main demand	Divided attention, task switching, and resource allocation.	Sustained readiness, signal detection, and response preparation.
Behavioral signature	Selective performance decline, particularly in monitoring-related components.	Maintained or improved response efficiency under some higher-frequency settings.
Ocular signature	Attentional narrowing toward central or high-priority components, accompanied by modest whole-stream ocular shifts.	Higher blink rate, gaze speed, and saccade proxy with lower valid-gaze fraction under high difficulty.
EEG signature	Hjorth features and frontal relative powers showed descriptive differences between serial and parallel execution.	Hjorth Mobility and Complexity showed descriptive modulation with event difficulty; frontal theta and beta increased slightly from low to high difficulty, while gamma remained elevated relative to baseline.
HMI implication	Protect peripheral cues during high continuous load.	Detect covert readiness cost during apparently quiet monitoring.

6. Discussion

The results support the view that pilot workload in automated environments should be interpreted in relation to task temporal structure. Continuous multi-task control and sparse event vigilance differ not only in demand level but also in demand type. Continuous multi-task control requires active resource allocation across concurrent subtasks, whereas sparse event vigilance demands sustained readiness during periods of limited overt action.

In the continuous multi-task control contrast, serial and parallel execution shared an identical subtask set, isolating the effect of concurrent execution. Table 2 and Figure 4 suggest that parallel execution produced a meaningful reduction in task performance. The decline varied across subtasks, with system monitoring exhibiting the largest descriptive reduction while communications and resource management remained relatively high. This pattern supports the interpretation that continuous demand prompts participants to prioritize salient or continuously actionable components at the expense of peripheral monitoring. The corresponding whole-stream ocular metrics shifted modestly, and frontal relative band powers shifted toward higher delta and beta with lower gamma. The principal continuous multi-task control interpretation therefore rests on the selective performance decline and AOI fixation observations, with EEG and Hjorth features treated as descriptive physiological context.

In the sparse event vigilance condition, Figure 6 shows that improved or maintained behavioral performance under higher event density can coexist with elevated workload. The accompanying eye-tracking changes in Figure 7 and Hjorth features in Figure 9 suggest that the participants' monitoring state evolved as task conditions changed. Because performance remained efficient, this pattern is more consistent with anticipatory readiness and arousal than with exhausted resources under concurrent task competition. Frontal relative gamma was higher during vigilance runs than at baseline, and theta and beta increased modestly from low to high difficulty, supplying secondary spectral context rather than the primary evidence for the sparse event vigilance effect. This dissociation between performance and physiology indicates that sparse event vigilance may impose covert cognitive costs that remain hidden if assessment relies solely on accuracy or response time.

For flight deck HMI design, these findings argue against treating workload as a single scalar state. An adaptive interface that detects overload only through degraded performance may fail to identify sustained anticipatory readiness during sparse monitoring. Future alerting and information-management strategies should therefore consider both overt performance and physiological indicators of readiness, arousal, and attentional allocation.

6.1 Limitations

Several limitations should be considered when interpreting these results. First, condition-order metadata, eye-tracker sampling and calibration information, full EEG preprocessing parameters, and final trial-level statistical models were not available for all analyses. Second, the MATB-II performance table is based on platform-derived aggregate scores; participant-level subtask data would enable stronger estimates of individual variability. Third, the paired eye-tracking analyses retained 10 participants, and EEG relative-power summaries should be interpreted descriptively, limiting the precision of band-power estimates. Fourth, although the simulator-based vigilance paradigm is more operationally grounded than a simple laboratory vigilance task, it remains an abstraction of real flight operations. Fifth, multimodal physiological recording can introduce missing data and artifact sensitivity, particularly during joystick operation and head movement. Finally, these results should be interpreted as evidence regarding task-structure-dependent workload profiles rather than a finalized operational workload classifier.

6.2 Implications for Workload-Adaptive HMI Design

The contrast between continuous multi-task control and sparse event vigilance yields several design implications for future flight decks. Adaptive systems should distinguish continuous overload from sparse readiness demand. Under high continuous workload, interfaces should preserve the salience of secondary monitoring information so that peripheral system states are not lost when attention narrows toward central control tasks. Under sparse event vigilance, alerting systems should not assume that low overt activity implies low cognitive demand; instead, monitoring support should account for readiness, expectation, and possible fatigue during extended low-event periods.

These findings also support multimodal workload assessment. Behavioral performance remains essential because it directly reflects task success. However, behavioral data alone may overlook compensatory effort and covert readiness cost. Ocular and EEG measures can identify when a pilot is sustaining performance through increased internal effort, or when a monitoring state is physiologically distinct from simple rest. For adaptive HMI design, the practical implication is that workload state should be classified by demand type and operational context, not by magnitude alone.

7. Conclusion

This study suggests that task temporal structure is an important determinant of pilot workload in highly automated environments. Continuous multi-task control and sparse event vigilance were associated with distinct workload profiles. Continuous multi-task control was characterized by selective performance degradation, attentional narrowing, and sustained resource utilization. Sparse event vigilance shifted demand toward sustained anticipatory readiness; preserved or slightly improved performance, together with ocular changes and descriptive Hjorth patterns, is more consistent with anticipatory readiness and arousal than with the resource-depletion pattern often associated with multitasking. The frontal relative delta, theta, beta, and gamma analyses provide secondary spectral context within this evidence set.

These findings suggest that workload assessment and adaptive HMI design should move beyond a single workload scale. In future automated flight decks, the cognitive cost of prolonged readiness during low-activity monitoring periods should be considered explicitly alongside conventional overload indicators.

8. Contact Author Email Address

The contact author email address is mailto: xiaowenduo@sjtu.edu.cn.

9. Copyright Statement

The authors confirm that they, and/or their company or organization, hold copyright on all of the original material included in this paper. The authors also confirm that they have obtained permission, from the copyright holder of any third party material included in this paper, to publish it as part of their paper. The authors confirm that they give permission, or have obtained permission from the copyright holder of this paper, for the publication and distribution of this paper as part of the ICAS proceedings or as individual off-prints from the proceedings.

References

[1] N. B. Sarter, D. D. Woods, and C. E. Billings. Automation surprises. In G. Salvendy, editor, *Handbook of Human Factors and Ergonomics*, pages 1926–1943. Wiley, 1997.

[2] R. Parasuraman, R. Molloy, and I. L. Singh. Performance consequences of automation-induced complacency. *The International Journal of Aviation Psychology*, 3(1):1–23, 1993.

[3] J. S. Warm, R. Parasuraman, and G. Matthews. Vigilance requires hard mental work and is stressful. *Human Factors*, 50(3):433–441, 2008.

[4] S. M. Casner and J. W. Schooler. Vigilance impossible: Diligence, distraction, and daydreaming all lead to failures in a practical monitoring task. *Consciousness and Cognition*, 35:33–41, 2015.

[5] A. R. Kyle, B. Rouser, R. C. Paul, and K. A. Jurewicz. Quantifying pilot performance and mental workload in modern aviation systems: A scoping literature review. *Aerospace*, 12(7):626, 2025.

[6] G. Luzzani, I. Buraioli, D. Demarchi, and G. Guglieri. A review of physiological measures for mental workload assessment in aviation: A state-of-the-art review of mental workload physiological assessment methods in human-machine interaction analysis. *The Aeronautical Journal*, 128(1323):928–949, 2024.

[7] J. Cegarra, B. Valery, E. Avril, C. Calmettes, and J. Navarro. Openmatb: A multi-attribute task battery promoting task customization, software extensibility and experiment replicability. *Behavior Research Methods*, 52:1980–1990, 2020.

[8] R. M. Touchstone. The effect of increased monitoring load on vigilance performance. Technical Report AM77-18, FAA Office of Aviation Medicine, 1977.

[9] S. Chikhi, N. Matton, and S. Blanchet. Eeg power spectral measures of cognitive workload: A meta-analysis. *Psychophysiology*, 59(6):e14009, 2022.

[10] A. Hag, F. Al-Shargie, D. Handayani, and H. Asadi. Mental stress classification based on selected electroencephalography channels using correlation coefficient of hjorth parameters. *Brain Sciences*, 13(9):1340, 2023.