

## Article

# The Impact of Various Cockpit Display Interfaces on Novice Pilots' Mental Workload and Situational Awareness: A Comparative Study

Huimin Tang <sup>1</sup>, Boon Giin Lee <sup>1,2</sup>, Dave Towey <sup>1</sup> and Matthew Pike <sup>1,\*</sup>

<sup>1</sup> School of Computer Science, University of Nottingham Ningbo China, 199 Taikang East Road, Ningbo 315100, China; huimin.tang@nottingham.edu.cn (H.T.); boon-giin.lee@nottingham.edu.cn (B.G.L.); dave.towey@nottingham.edu.cn (D.T.)

<sup>2</sup> Nottingham Ningbo China Beacons of Excellence Research and Innovation Institute, Ningbo 315101, China

\* Correspondence: matthew.pike@nottingham.edu.cn

**Abstract:** Future airspace is expected to become more congested with additional in-service cargo and commercial flights. Pilots will face additional burdens in such an environment, given the increasing number of factors that they must simultaneously consider while completing their work activities. Therefore, care and attention must be paid to the mental workload (MWL) experienced by operating pilots. If left unaddressed, a state of mental overload could affect the pilot's ability to complete his or her work activities in a safe and correct manner. This study examines the impact of two different cockpit display interfaces (CDIs), the Steam Gauge panel and the G1000 Glass panel, on novice pilots' MWL and situational awareness (SA) in a flight simulator-based setting. A combination of objective (EEG and HRV) and subjective (NASA-TLX) assessments is used to assess novice pilots' cognitive states during this study. Our results indicate that the gauge design of the CDI affects novice pilots' SA and MWL, with the G1000 Glass panel being more effective in reducing the MWL and improving SA compared with the Steam Gauge panel. The results of this study have implications for the design of future flight deck interfaces and the training of future pilots.



**Citation:** Tang, H.; Lee, B.G.; Towey, D.; Pike, M. The Impact of Various Cockpit Display Interfaces on Novice Pilots' Mental Workload and Situational Awareness: A Comparative Study. *Sensors* **2024**, *24*, 2835. <https://doi.org/10.3390/s24092835>

Academic Editor: Wataru Sato

Received: 26 March 2024

Revised: 21 April 2024

Accepted: 26 April 2024

Published: 29 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The number of in-service airplanes has increased in the past 20 years, according to data from GAMA (<https://gama.aero/wp-content/uploads/2021ShipmentReport-2022-05-16-1.pdf>) (accessed on 6 July 2022). It follows that future air traffic will be more congested, increasing the number of factors that could lead to accidents (<https://www.iata.org/en/iata-repository/publications/economic-reports/global-outlook-for-air-transport---december-2023---report/>) (accessed on 19 April 2024)). In addition, increased congestion will lead to pilots having to process more information from air traffic controllers, which will increase the number of parallel tasks they need to perform [1]. Mental workload (MWL) is a psychological construct modeled on real activity [2] to describe how much mental effort a given task requires at a given time [3]. The majority of recent aviation accidents have seen high MWL levels [4], highlighting the importance of pilots' MWL in Flight Operation Safety Assessment (FOSA). Situational awareness (SA) [5] is another important factor in ensuring flight safety [6]. It represents operators' understanding of a specific situation and can be viewed as a three-step cognitive concept comprised of (1) *perception*, (2) *comprehension* and (3) *projection* [2]. SA-friendly systems improve pilot efficiency and increase flight safety [7].

The airplane's cockpit display interface (CDI) is one of the most important devices for pilot–airplane interaction [8]. The CDI provides critical airplane state information, such as the heading, altitude and speed. Endsley [5] pointed out that the CDI needs to support

pilots in building sufficient SA to maximize flight safety. Existing studies have shown the importance of interface design on operators' MWL [9]. However, current studies mainly focus on how to design the CDI to better support pilots' SA [10] or focus on how individual parts of the CDI, such as the navigation system, influence pilots' MWL [11,12], and rarely is the relationship between a CDI's support for SA and pilots' MWL considered.

Two research questions were proposed for this study to study how different subparts of the CDI influence pilots' SA and how the impact on SA influences their MWL:

RQ1: Do different designs of CDIs that provide the same information have different effects on a pilot's SA?

RQ2: What impact do different designs of CDIs with the same information have on pilots' MWL?

This study focuses on a specific combination of several displays on a CDI can provide pilots with spatial awareness to study the proposed RQs. Spatial awareness has been noted as one of the most important components of SA for aviation [13], which can be provided by three orientation variables—pitch, roll and heading—and three position variables: altitude, lateral path deviation and flight path location. In this study, the impacts of the selected displays' layouts on two specific Cessna-172 series CDI types, the Steam Gauge (SG) panel and the G1000 Glass (GG) panel, on novice pilots' SA and MWL were compared with each other using objective and subjective measures. The GG panel is a digital flight display that integrates all key flight data into a single display, while the SG panel displays information using gyroscopes. The experiment asked the participants to complete a flight task as a primary task and finish a parallel subtask to simulate a real-world flight scenario. The 4D multiple resources model [14] and the resources–performance relationship [15] were adopted to interpret the impact of the CDI design on participants' MWL and SA. Our study reaffirms the previous finding that the CDI design influences pilots' SA [8,16,17] and reveals how differently designed CDIs will influence pilots' MWL. Additionally, this work provides a further demonstration of applying MWL measures in the evaluation of real-world flight displays. Future flight displays can be informed by the results of this study, and CDI designers can understand the impact of different display designs on pilots' SA and MWL using the methodology employed in this study.

## 2. Related Work

Two aspects will be discussed in this section: (1) situational awareness (SA) and mental workload (MWL) and (2) information processing models.

### 2.1. Situational Awareness and Mental Workload

SA is the assessment of operators' correctness in comprehending a given situation [2]. It has been modeled as a three-step process: (1) perceiving the elements in an environment during a period (*perception*), (2) comprehending their meaning (*comprehension*) and (3) projecting their following status based on the understanding of the current situation (*projection*) [5]. Existing studies have identified a relationship between the CDI design and pilots' SA [8,16,17]. During a computer-based flight simulation, Andre et al. [16] presented navigation displays in three display formats and two color-coding types (six trials). They investigated the influence of an inside-out pilot-centered navigation display, an outside-in world-centered navigation display and a 3D outside-in navigation system on novice pilots' SA. The navigation displays were shown both in monochrome and color. Participants were instructed to accumulate as many points as possible in each of the six trials. The study found that the outside-in navigation display had a more positive effect on pilots' SA compared with the other display formats, and the use of color in displays improved situational awareness, particularly during emergency situations. Wei et al. [8] evaluated participants' SA levels for three differently designed GG panels by having them monitor an aircraft's instrument panel for 12 min, during which a complete flight was simulated (take-off, climb, cruise, descent and landing). The instrument panel was programmed to freeze at preset intervals during the task, and the participants were queried on their awareness of

the aircraft's current status through a series of questions. They found that the CDI design had a significant impact on pilots' SA, especially during the *projection* and *comprehension* stages. The SA under Boeing's CDI was significantly higher than that under Airbus's at the *projection* stage and significantly lower at the *comprehension* stage. Their findings suggested that although Boeing's CDI helped participants perceive more information, it was more challenging for them to combine the gathered information for understanding.

MWL can be used in evaluating the performance of the operator and the system [3] to improve the system design. Longo [18] emphasized that excessively high or low MWL would negatively influence human performance. Studies have demonstrated a similar relationship between flight performance and MWL [19–21]. Morris and Leung [19] asked different groups of novice pilots to finish different numbers of parallel tasks on a computer-based simulator that simulated a real-world flight task's different demand levels. The primary task required the participants to keep two crosshairs in the center of the screen using a joystick. The study found a decline in the participants' primary task performance under a higher MWL.

As reported by Wei et al. [20], the accuracy of manipulating the flight simulator decreased, while the reaction time increased as the pilots' MWL increased. A limited number of studies have examined the relationship between CDI design and pilots' MWL, with a focus on the navigation system (primary flight display) on the CDI. For example, Beringer and Ball [12] and Uenking and Hughes [22] investigated the optimal placement of the primary flight display, and Uenking and Hughes [22] and Davis et al. [23] examined ways to design it to minimize MWL. Both Davis et al. [23] and Uenking and Hughes [22] found that the new primary flight displays resulted in a lower MWL compared with conventional round dial instrumentation. Beringer and Ball [12] suggested that the field of view used for the primary flight display should not be less than 40 degrees. However, few studies have examined the effect of the entire CDI design on a pilot's MWL. Casner [11] compared the impact of the CDI with round dial and electronic instruments on pilots' MWL, but the small sample size precluded clear conclusions.

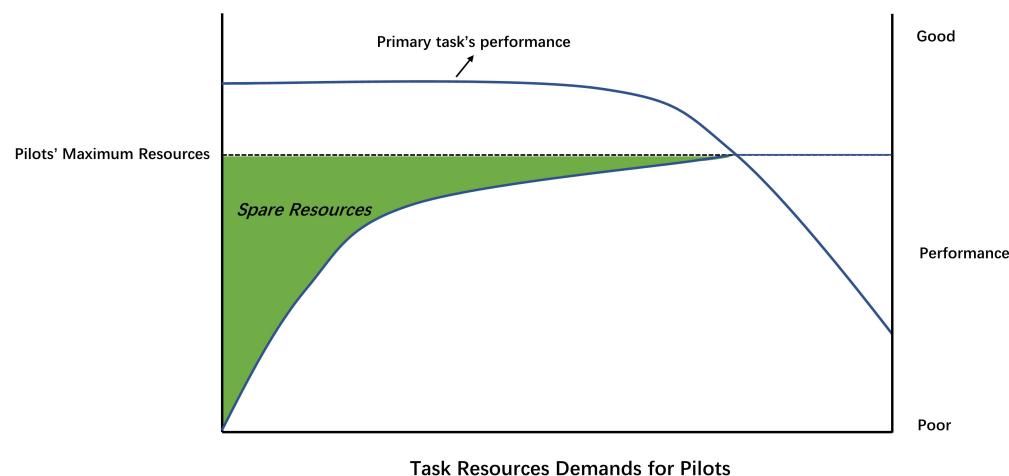
Studies have highlighted the interaction between SA and MWL and its effect on pilots' performance [8,24]. Lin and Lu [24] designed a simulated helicopter rescue task. The task difficulty was induced by different emergencies, such as some suddenly broken instruments or random alarm lights on the panel. They found that low SA led to a high MWL, and decreased pilots' performance, especially for unskilled pilots.

## 2.2. Information Processing Models

The 4D multiple resources model (MRM) proposed by Wickens [25] describes three stages of information processing: *perception*, *cognition* and *responding*. Different dimensions, such as the *visual* and *auditory* dimensions, might be involved in information processing, depending on the task demands, and consume the corresponding types of cognitive resources. This model is especially suitable for predicting probable inter-task interference during multitasking [25], meaning designers can predict such interference in advance [26] and identify resources demand conflicts [27]. Existing studies have also established that this model can accurately represent changes in MWL during multitasking [28]. When the tasks require resources from the same dimension, such as several tasks requiring *audio* resources to handle different sounds, the operator's MWL will be greater relative to task demands from differing resource dimensions. Resource competition happens frequently during piloting [14]. For example, the competition for *visual* resources is intense due to the need to process information from a large number of instruments [29] or the requirements when simultaneously perceiving information from both the instruments and the outside environment [30]. The MRM has been widely used in exploring the relationship between the pilots' workload and their performance [31]. Thorpe et al. [31] used the MRM to explain their findings that the negative effect of two parallel tasks requiring different resources (visual aviation task and auditory secondary task) was less than that of two parallel tasks

requiring the same resources (visual aviation task and visual secondary task) on pilots' flight performance.

The resources–performance relationship adapted by Sharples and Megaw [15] (p. 520) showed that the human capacity for processing information is limited (Figure 1). When the spare capacity approaches zero, primary task performance will decline. Mental overload occurs if the required resources exceed the available resources ([15], p. 543). Combined with the MRM [25], when several tasks (including the primary task) require resources from the same dimension, the remaining capacity in this dimension declines, increasing the possibility of primary task performance degradation. Therefore, resource competition in one dimension could induce both a higher MWL and worse performance. This has been shown in existing pilot-related studies [8,32].



**Figure 1.** Resources–performance relationship, adapted from Sharples and Megaw [15] (p. 520).

### 3. Methods

#### 3.1. Hypotheses

Pilots will consume fewer cognitive resources on building up spatial awareness if the displays related to these variables are appropriately constructed. Wickens [13] stated that maintaining SA can be viewed as a cognitive task that requires cognitive resources to process selective information and subsequently may impact pilots' MWL. The Cessna 172 SG panel (Figure 2) and the Cessna 172s GG panel (Figure 3) were chosen for this study because of their widespread use and distinctive differences in spatial awareness-related instrumentation design, including the airspeed indicator, altimeter, attitude indicator and heading indicator ((a–d) in Figures 2 and 3).

The following hypotheses were developed based on the research questions:

- **HPSA** (Operator's SA)

Under the same task difficulty, there will be a significant difference in SA between two different CDIs.

- **HPMWL** (Operator's MWL)

For the same task, the operator's MWL will be significantly lower with a CDI that has a higher SA level than with a CDI that has a lower SA level.

**HPSA** is proposed to address RQ1 and is based on the findings of Wei et al. [8], who identified a relationship between different CDI designs and their impact upon pilots' SA (discussed in Section 2.1). In addition, Endsley [33] pointed out that the system interfaces that were designed based on SA-oriented interface design principles [33] have been shown to better support operators' SA, suggesting the interface design's impact on the operator's SA. Compared with the digital display, the pointer-type ones used on the SG panel have a higher possibility of misreading, especially when the operator is asked to make quantitative readings from the display [34]. Such displays have a higher possibility of depleting the operator's resources for maintaining sufficient SA.



**Figure 2.** The Cessna 172 Steam Gauge (SG) panel.



**Figure 3.** The Cessna G1000 Glass (GG) panel.

HPMWL was drawn from the resources–performance relationship ([15], p. 520) and MRM [25] for exploring RQ2. Facing limited resources, operators will theoretically have more spare capacity for other parallel cognitive tasks and have less intense competition in resources when SA can be maintained from the CDI using fewer resources, especially when the parallel tasks require resources from the same dimension as maintaining SA. This decreases the possibility of a high MWL and mental overload. A lower MWL may result from the CDI that better supports the pilot’s SA.

### 3.2. Task

#### 3.2.1. Primary Task

A flight simulator running *Microsoft Flight Simulator 2020* was provided by the School of Aerospace at the University of Nottingham Ningbo China (UNNC) and was used to conduct the primary task. Participants were required to pilot the airplane, maintaining a heading of 90° E at an altitude of 2500 feet (0.76 km) for 4 minutes. The task was completed at nighttime in the simulation environment, meaning that there were little-to-no visual landmark cues for participants to utilize. Instead, the participants had to rely solely on the readings from the CDI (Figure 4) to complete the task.

There were several considerations in the primary task design to reduce the influence of factors unrelated to the CDI. The task tried to reduce extra operations from the task that might distract participants, such as communicating with the traffic controller. Because this study focused on the instruments that provide spatial awareness of the CDI, the tasks were designed such that the participants were required to regularly consult the CDI.



**Figure 4.** Flight simulator and environment for primary task.

### 3.2.2. Secondary Task

The N-back task [35] was the secondary task, which required the participants to speak out the number that previously appeared  $n$  items ago in the task. At the **perception** stage, referring to the MRM [25], both the primary and N-back tasks needed to use **visual** resources, making the competition in the *visual* dimension become intense and resulting in a significant increase in the participants' MWL. Based on the resources–performance relationship ([15], p. 520), the use of *visual* resources by the N-back task resulted in a decrease or even a complete consumption of spare resources for the primary task, leading to a degradation in the primary task's performance. Based on the resources–performance relationship ([15], p. 520), the utilization of *visual* resources by the N-back task led to a partial or total depletion of spare resources for the primary task, resulting in a decline in primary task performance. At the **cognition** stage, the competition for *verbal* resources between the N-back task and the primary task became more intense as the N-back task's difficulty increased. This implies that as the difficulty of the N-back task increases, primary task performance may degrade, and mental workload (MWL) may increase. Hence, we expected the following:

- As the difficulty of the N-back task increased, the performance of the primary task degraded.
- As the difficulty of the N-back task increased, the MWL of the participants rose.

Participants were required to complete three rounds of tasks with each CDI:

- Single flight;
- Flight + one-back task;
- Flight + two-back task.

A “baseline” task of sitting still for one minute was conducted at the beginning of the study to record the participants' physiological data that could reflect their MWL (discussed further below) during a resting state. The order of the six tasks for each participant was arranged using Latin square rotation to minimize the effects of the task order on the results [36] (see Table 1). The number of participants was larger than six. Thus, the task order started from the beginning of the table again after one round.

**Table 1.** Experiment order.

| Participant | Order of the Experiment |                        |                        |                        |                        |                        |
|-------------|-------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
|             | 1                       | 2                      | 3                      | 4                      | 5                      | 6                      |
| 1           | single-flight digital   | 1 back digital         | 2 back analogue        | 2 back digital         | 1 back analogue        | single-flight analogue |
| 2           | 1 back digital          | 2 back digital         | single-flight digital  | single-flight analogue | 2 back analogue        | 1 back analogue        |
| 3           | 2 back digital          | single-flight analogue | 1 back digital         | 1 back analogue        | single-flight digital  | 2 back analogue        |
| 4           | single-flight analogue  | 1 back analogue        | 2 back digital         | 2 back analogue        | 1 back digital         | single-flight digital  |
| 5           | 1 back analogue         | 2 back analogue        | single-flight analogue | single-flight digital  | 2 back digital         | 1 back digital         |
| 6           | 2 back analogue         | single-flight digital  | 1 back analogue        | 1 back digital         | single-flight analogue | 2 back digital         |

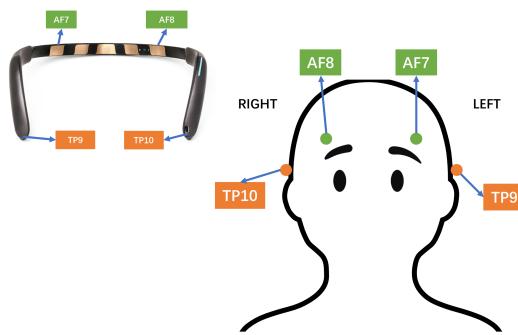
### 3.3. Participants

Ten participants (5 male and 5 female) with an average age of 21 years were recruited to take part in the experiment. The participants were all from UNNC. All participants had normal or corrected vision and reported no history of head trauma or brain damage. All the participants had less than 20 h of flight experience on flight simulators but no experience with real-world piloting to ensure there would not be confounding variables caused by different experience levels or different educational backgrounds. The study was approved by the faculty's ethics committee, and all participants provided informed consent prior to beginning the study.

### 3.4. Data Collection

*PsychoPy* (<https://www.psychopy.org/>) (accessed on 21 April 2021) was used to present the N-back task and record associated performance data as well as to stream sensor data using the lab streaming layer (LSL) protocol (<https://github.com/sccn/labstreaminglayer> (accessed on 21 April 2021)). *LabRecorder* (<https://github.com/labstreaminglayer/App-LabRecorder> (accessed on 21 April 2021)) was used to receive the streams transferred via the LSL and record them together into an XDF file with time synchronization between them onto the researcher's laptop as the study was being completed.

Changes in MWL have been shown to be reflected in changes in electroencephalography (EEG) components. Among them, the  $\theta$  band (4~7 Hz) and  $\alpha$  band (8~12 Hz) [37] have been successfully used in many pilots' MWL-related experiments (e.g., [38]). Several studies have found that the  $\alpha$  band decreased in the frontal and central regions of the cerebral cortex [39] and parietal regions [40], whereas the  $\theta$  band increased as the MWL increased, especially in the frontal regions [41]. *Muse2* (<https://choosemuse.com/muse-2/> (accessed on 21 April 2021)), a noninvasive EEG headband with four sensors on the forehead and ears, and a sampling rate of 256 Hz were used in this study. EEG measures electrical activity originating from different brain regions by placing sensors over the scalp [37]. The headband sits on the subject's forehead (locations AF7 and AF8) and rests on the ears (TP9 and TP10), as shown in Figure 5, using the international 10-20 EEG placement system [42]. Several studies have raised concerns regarding the accuracy and correctness of *Muse2*'s readings (e.g., [43]), but the device remains popular because it is highly portable, minimally invasive, easy to use and low in cost, especially when compared with clinical EEG systems [43]. *Muse2* has been proven to be sensitive in monitoring users' brain activities [44]. *BlueMuse* (<https://github.com/sccn/xdf/releases/tag/v1.12> (accessed on 21 April 2021)) was used to collect data from *Muse2* and to interface with the LSL.



**Figure 5.** *Muse2* sensor locations using the 10-20 EEG placement system.

Heart rate variability (HRV) is another common physiological measure and has a known relationship with MWL [45]. The results of spectral analysis on HRV show that the low-frequency band (LF, 0.04~0.15 Hz) represents cardiac sympathetic nervous activities, which are more active under heavy workload and stress, while the high-frequency band (HF, 0.15~0.40 Hz) relates to the cardiac vagal nervous activities, which are in charge of cardiac activities in normal cases [37]. The LF/HF ratio, representing the balance of the sympathetic tone and vagal tone, can be used as a reliable index to reflect the changes in MWL [46].

According to Murata [47], the LF/HF ratio increases as the MWL increases. A *Polar H10* (<https://www.polar.com/uk-en/sensors/h10-heart-rate-sensor> (accessed on 21 April 2021)), a sports chest strap, was used to capture the participants' HRV during the study. The H10 is an ECG device that has been validated in previous studies measuring HRV under naturalistic task conditions [48]. *BLEPolarDirect* (<https://github.com/roberttwomey/nosc-toolkit> (accessed on 21 April 2021)) was used to pass the data returned (via Bluetooth) from the *Polar H10* to *PsychoPy* via the LSL.

### 3.5. Procedure

Before the start of the study, the participants were given 20 minutes to familiarize themselves with the flight simulator. At the beginning of the study, the participants were asked to sit still and relax for one minute while their physiological data were recorded using *Muse2* and the *Polar H10*. The baseline data were used to ensure the validity of the subsequent data collected and to eliminate noise in the EEG data. At the beginning of each task in the study, the airplane in the flight simulator was set to a starting altitude of 2500 feet (0.76 km) and an airspeed of 163 knots, heading in the direction of 90 degrees east. This airspeed is typical for a Cessna 172 series airplane [49]. The participants were expected to maintain the same altitude and direction during the experiment (Figure 6). Maintaining the airspeed requires pilots to consider various factors such as outside weather conditions, wind speed, air friction and the airplane's engine power. Although the weather and wind speed were constant in this study, changes in altitude have an impact on air drag and engine power and vice versa, resulting in changes in airspeed. As the participants in the study were novice pilots and may not have had the ability to maintain both the airspeed and altitude, they were not required to control the airspeed during the experiment.



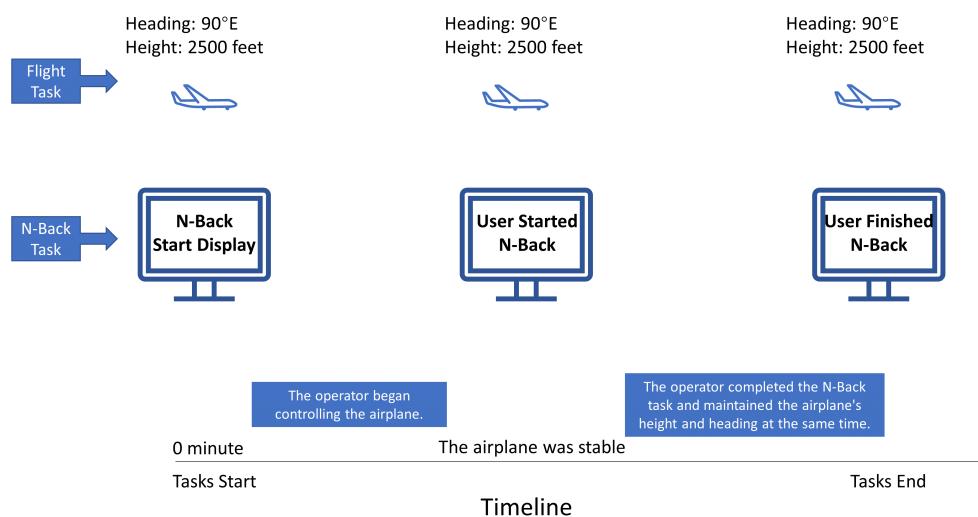
(a) During study 1  
(b) During study 2

**Figure 6.** During study.

The participants were required to perform an N-back task concurrently with a primary flight task. The N-back task was initiated at the same time as the flight task, and the numbers for the task were displayed on a separate screen as the participant started to operate the flight simulator. There was no set time for when the participants should start the N-back task, and they were able to begin it whenever they deemed their airplane to be stable. The participants were asked to verbally provide the answer to the N-back task. When they completed the N-back task, the flight task ended (Figure 7). The N-back answers, EEG data and HRV data were recorded using *PsychoPy*, while the flight performance data were recorded by the flight simulator.

After each task, the participants completed the NASA Task Load Index (NASA-TLX) assessment to evaluate their experience during the task. The NASA-TLX is a commonly used subjective MWL measure [50]. It has been shown to be quite reliable [51] and has been used extensively in aviation contexts [1,32]. The NASA-TLX predicts MWL by asking the participants to subjectively evaluate six linear scales: mental demand, physical demand, temporal demand, performance, effort and frustration.

After completing the assessment, the participants took one minute of relaxation time before beginning the next task. This process was repeated for all subsequent tasks.



**Figure 7.** Procedure.

#### 4. Results

In this study, all of the dependent variables (flight performance, NASA-TLX score, HRV and EEG) were determined to be non-parametric data based on the normal distribution of their values. All data were determined to be distribution-free using a Shapiro–Wilk test ( $p < 0.05$ ). All dependent variables were either ordinal (NASA-TLX score) or continuous (flight performance, HRV and EEG) data. Therefore, the Wilcoxon test was used to examine the impact of the type of CDI (two levels) on the dependent variables, while the Friedman test was applied to address the influence of the difficulty of the N-back task (three levels) on the dependent variables [52]. Median rather than mean values were used in the data analysis because of the relatively small sample size and to reduce outlier effects [53].

##### 4.1. Flight Performance

The participants were required to maintain the airplane's altitude and heading. Consequently, the flight performance criteria were based on the time the operators kept the airplane at different altitudes and headings during the task (Table 2).

**Table 2.** Flight performance criteria.

|                    | Heading (°)    | Altitude (Feet)      |
|--------------------|----------------|----------------------|
| Perfect (5 points) | $90 \pm 10$    | $2500 \pm 100$       |
| Good (3 points)    | 65–80, 100–115 | 2250–2400, 2600–2750 |
| Normal (2 points)  | 50–65, 115–130 | 2100–2250, 2750–2900 |
| Poor (0 point)     | Others         | Others               |

After applying the scoring criteria, a final score for each participant was calculated as follows:

$$\text{scoreList} = [5, 3, 2, 0] \quad (1)$$

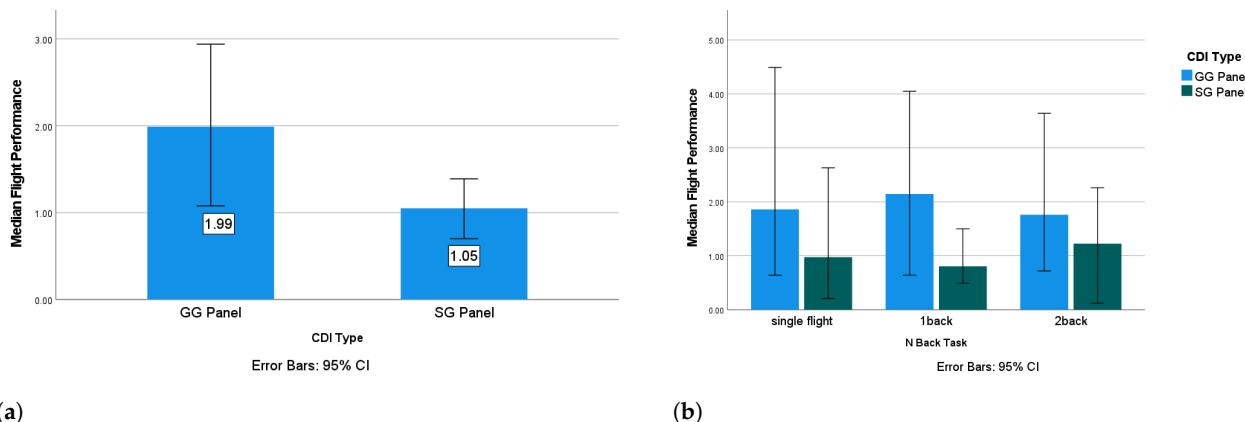
$$\text{performanceList} = [\text{Perfect}, \text{Good}, \text{Normal}, \text{Poor}] \quad (2)$$

$$\text{rating} = \frac{\text{time of keeping at this performance level}}{\text{total flight time}} \quad (3)$$

$$\text{finalScore} = \sum_{i=1}^4 \text{rating}_{\text{performanceList}[i]} \times \text{scoreList}[i] \quad (4)$$

The flight task required the operator to have sufficient SA about the airplane because its altitude and heading would continually change based on real-time spatial awareness. For example, the operators needed to understand that the current data on the CDI indicated a “downtrend”, and they needed to make the right decision based on the information on the CDI to keep the altitude from changing. Therefore, flight performance acted as the mapping of the SA level in this study.

The flight performance with the GG panel was significantly better than on that with the SG panel ( $Z = -3.816, p < 0.001$ ), which supports the hypothesis that the CDI type would have an impact on flight performance (HPSA) and was consistent with previous findings [8,16,17]. The overall better flight performance with the GG panel (Figure 8a) suggests that the GG panel is the more SA-friendly CDI compared with the SG panel. No significant impact from the difficulty of the N-back task on flight performance was detected. To further examine the impact of the CDI type on flight performance under different N-back conditions, we conducted follow-up Wilcoxon tests on flight performance grouped by CDI type under each of the three N-back conditions. The results revealed significantly better flight performance with the GG panel under the single-flight ( $Z = -2.395, p = 0.017$ ) and one-back ( $Z = -2.497, p = 0.013$ ) conditions, but there was no significant difference between the CDI types under the two-back condition. This suggests that when the multi-task condition does not exceed the operator’s limited cognitive resources, the interface design has a significant influence on their SA level. However, when participants experience an extremely high MWL or mental overload (such as in the two-back condition in this study), they may make task management errors [54], causing the flight task to become their secondary task and resulting in a lack of cognitive resources. This may explain the non-significant difference in SA level between the two CDI types under the two-back condition. Despite this, the median scores still showed that the SA level was higher when using the GG panel compared with the SG panel under the two-back condition (Figure 8b), suggesting that the GG panel is the more SA-friendly CDI.



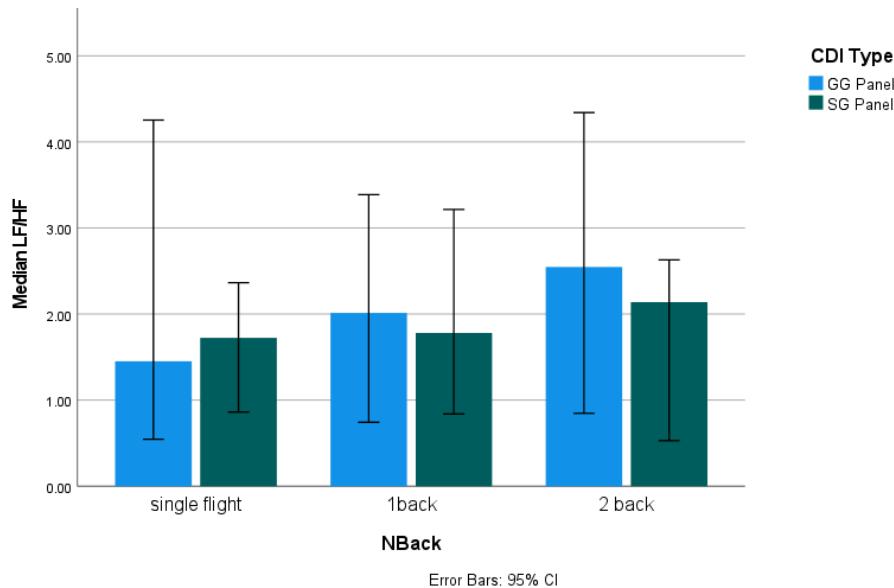
**Figure 8.** Median flight score under different conditions. (a) Median flight scores for different CDIs; (b) Median flight scores grouped by CDI and N-back conditions.

#### 4.2. MWL Measures

##### 4.2.1. HRV Data

The *hrvanalysis* library in Python [55] was used to calculate the LF/HF ratio. No significant differences in the LF/HF ratio were found between the different CDI types or N-back conditions. There was also no clear trend in the HRV data between the different conditions (Figure 9). This may be because the flight task in this study was designed to be a “cruise” stage, which is typically associated with relatively low cognitive demands. According to Wilson [21], HRV data tend to have more noticeable changes during the takeoff and final landing stages of a flight, while the HRV data during other stages of the flight may only exhibit slight changes. HRV may not be sensitive enough to identify

differences in MWL under different task demands if the demand changes are not large enough [56].



**Figure 9.** Median LF/HF ratio under different conditions.

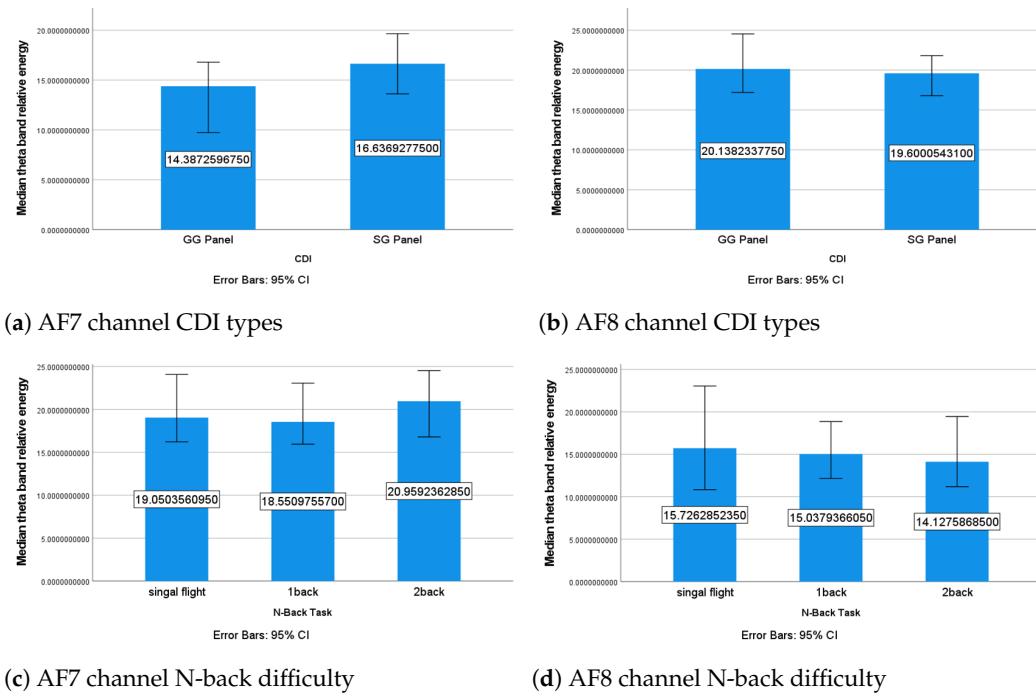
#### 4.2.2. EEG Data

The EEG data were first channel re-located based on the international 10–20 system using *EEGLab* (<https://github.com/sccn/eeglab> (accessed on 21 April 2021)) to ensure the channel names in the data file had the correct correspondence with the locations on the scalp. Then, the signals were re-referenced using EEGLab to ensure the signals in each channel were not impacted by the different distances between the sensors and the reference sensor. The mean mastoids (MMs) of TP9 and TP10 were used as the re-referencing point, which is a widely used re-referencing method [57]. Then, after filtering the EEG data using a 1–30 Hz filter, a wavelet transform [58] was applied to extract the  $\alpha$  band and  $\theta$  band required for this study from the original EEG signal, which is a method for decomposing the EEG signal into its frequency components through dilatations, contractions and shifts [59]. Additionally, the wavelet transform is suitable for the spectral analysis of each decomposed frequency band, making it easier to calculate the relative energy of the bands.

##### The $\theta$ Band

The statistical analysis revealed that the CDI type and N-back task did not have a significant effect on the relative energy of the  $\theta$  band in either the AF7 and AF8 channels. The overall relative energy of the  $\theta$  band was lower under the GG panel and highest under the two-back condition in the AF7 channel (Figure 10a,c). Previous research has shown that when individuals engage in mental activities that require higher cognitive resources, the  $\theta$  power tends to increase more over the frontal cortex area of the brain [60], which is the area that the AF7 and AF8 electrodes of *Muse2* correspond to. Therefore, the lower relative energy of the  $\theta$  band in the AF7 channel when operating the GG panel observed in this study may indicate that the operators had a lower MWL under this condition. This finding supports the conclusion from the flight performance analysis that the GG panel is the more SA-friendly CDI in this study, even though it does not fully support the hypothesis that the MWL would be lower under the GG panel (HPMWL). The higher relative energy of the  $\theta$  band in the AF7 channel under the two-back condition suggests that the two-back task induced a higher MWL than the single-flight and one-back conditions. This is consistent with the conclusion from the flight performance analysis that the non-significant difference between the CDI types under the two-back condition may have been caused by the participants experiencing extremely high MWLs or mental overload. However, the

findings in the AF8 channel showed a trend that was opposite to that in the AF7 channel (Figure 10b,d). The relative energy of the  $\theta$  band was higher for the GG panel, and the single flight induced the highest MWL among the three N-back conditions.



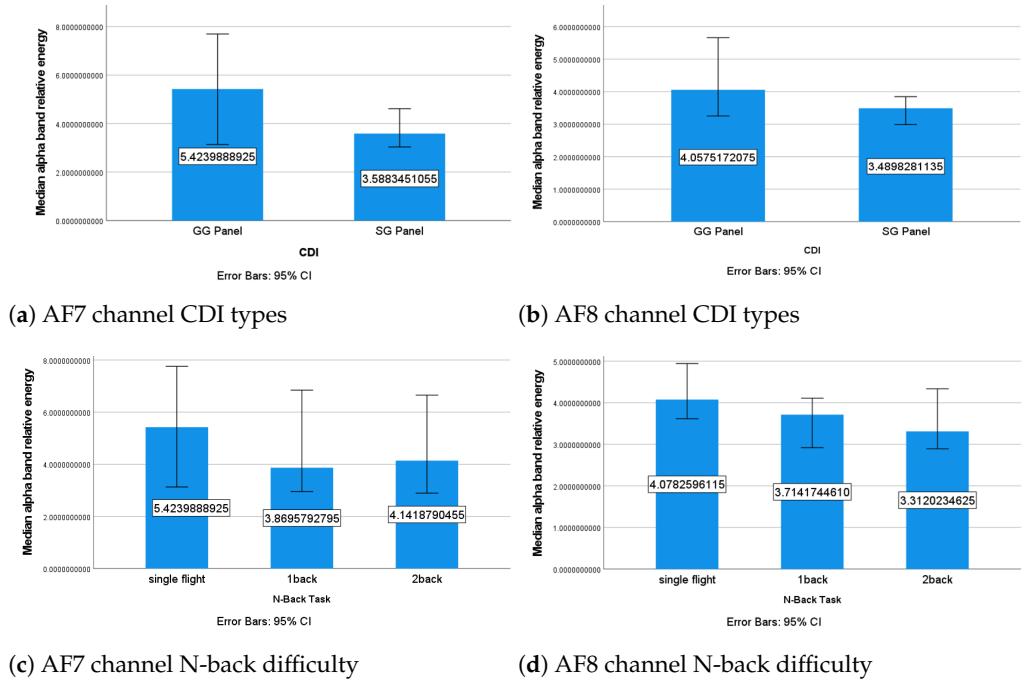
**Figure 10.** Median  $\theta$  band's relative energy under different conditions.

#### The $\alpha$ Band

When using the GG panel, the relative energy of the  $\alpha$  band was significantly higher in the AF8 channel ( $Z = -2.211, p = 0.027$ ) and marginally significantly higher in the AF7 channel ( $Z = -1.8, p = 0.072$ ) compared with the SG panel (Figure 11a,b). When the Wilcoxon test was applied to the relative energy of the  $\alpha$  band in the AF8 channel grouped by CDI type under each of the three N-back conditions, a significantly higher  $\alpha$  band relative energy for the GG panel was found under the single-flight condition ( $Z = -2.293, p = 0.022$ ), and a marginally significantly higher  $\alpha$  band relative energy for the GG panel was found under both the one-back ( $Z = -2.211, p = 0.062$ ) and two-back conditions ( $Z = -1.886, p = 0.059$ ).

The relative energy of the  $\alpha$  band showed a significant difference between the different N-back conditions in both the AF7 ( $\chi^2 = 6.4, p = 0.041$ ) and AF8 ( $\chi^2 = 6.3, p = 0.043$ ) channels. After applying a Bonferroni correction [61], the  $\alpha$  band's relative energy under the single-flight condition was significantly higher than under the one-back condition in the AF7 channel ( $p = 0.034$ ) and significantly higher than under the two-back condition in the AF8 channel ( $p = 0.018$ ) (Figure 11c,d). This suggests that the N-back tasks increased the MWL of the participants compared with the single-flight condition.

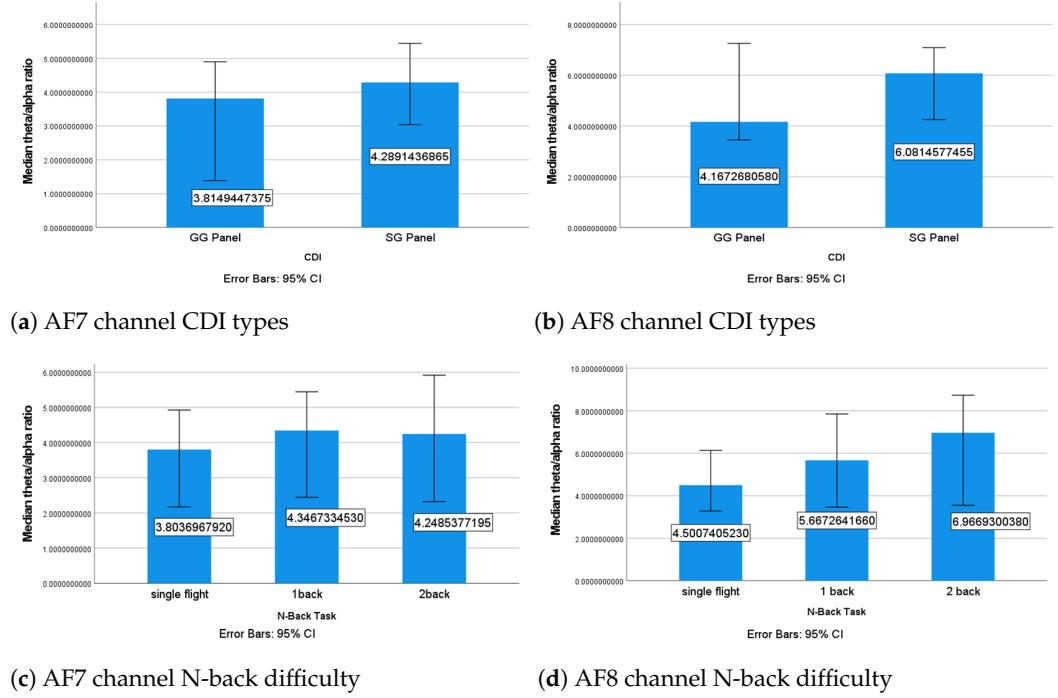
The results of this study suggest that the GG panel may be designed to be more spatial awareness-friendly, inducing a lower MWL compared with the SG panel under the same conditions. This was inferred from the decrease in the overall relative energy of the  $\alpha$  band in both the AF7 and AF8 channels while using the SG panel, as the  $\alpha$  waves' increment has been shown to reflect a relaxed mental state [62]. The N-back tasks also increased the MWL of the participants compared with the single-flight condition, as indicated by the significantly increased relative energy of the  $\alpha$  band in the AF7 and AF8 channels. It is worth noting that there were some marginally significant results ( $p < 0.075$ ), and increasing the sample size may further strengthen these results [63].



**Figure 11.** Median  $\alpha$  band's relative energy under different conditions.

#### The $\theta/\alpha$ Ratio

No significant impact from the CDI types on the  $\theta/\alpha$  ratio was identified in either channel. The overall  $\theta/\alpha$  ratio under the SG panel was higher in both channels (Figure 12a,b). Since the increase in MWL can be reflected by the increment of the  $\theta/\alpha$  ratio [64], we interpreted that the participants' MWL under the SG panel was descriptively higher than that under the GG panel, which could support the trend in HPMWL.



**Figure 12.** Median  $\theta/\alpha$  ratio under different conditions.

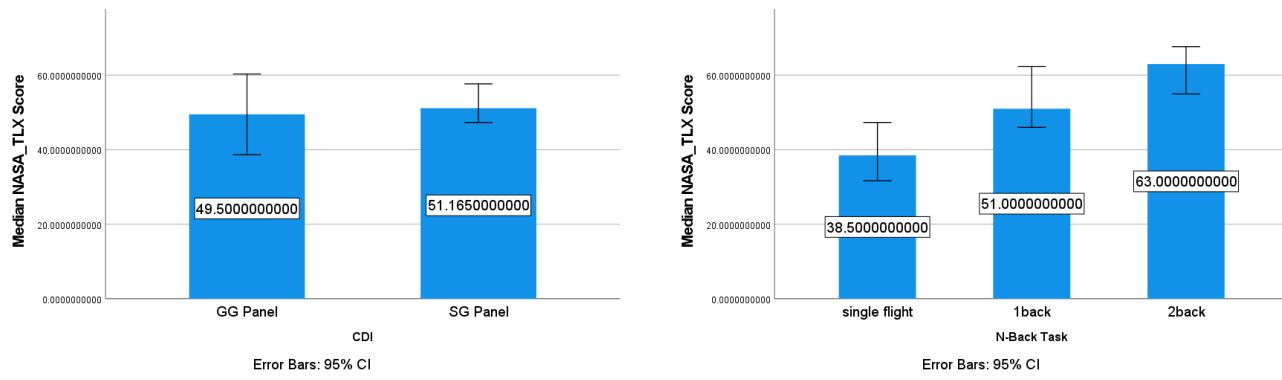
There was no significant difference in the  $\theta/\alpha$  ratios between the different CDI types in either the AF7 or AF8 channels. The overall  $\theta/\alpha$  ratio was higher under the SG panel in both channels (Figure 12a,b). Since an increase in MWL can be reflected by an increase in

the  $\theta/\alpha$  ratio [64], it can be concluded that the participants' MWL was descriptively higher while using the SG panel compared with the GG panel. This finding supports **HPMWL**.

There was no significant difference in the  $\theta/\alpha$  ratio between the different N-back conditions in the AF7 channel (Figure 12c), but a significant difference was found in the AF8 channel ( $\chi^2 = 9.1, p = 0.011$ ). After applying a Bonferroni correction [61], a significant difference was found between the single-flight and one-back conditions ( $p = 0.034$ ) and between the single-flight and two-back conditions ( $p = 0.022$ ) in the AF8 channel. The  $\theta/\alpha$  ratio under the single-flight condition was significantly lower than the one-back and two-back conditions (Figure 12d). This suggests that the N-back tasks increased the MWL of the participants during the flight task, which is consistent with the findings obtained from the analysis of the  $\alpha$  band's relative energy.

#### 4.2.3. NASA-TLX Score

The overall NASA-TLX score was used for analysis. It is worth noting that while there was no significant difference in NASA-TLX scores across the CDI types, there was a slightly higher median score when the participants used the SG panel. This suggests that the SG panel may have caused a slightly higher MWL compared with the GG panel, although the difference was not statistically significant. A significant impact of the N-back condition's difficulty on the NASA-TLX score was found ( $\chi^2 = 29.1, p < 0.001$ ). After applying a Bonferroni correction [61], the NASA-TLX score under the two-back condition was significantly higher than that under the single-flight ( $p < 0.001$ ) and one-back conditions ( $p < 0.001$ ). This suggests that the N-back tasks successfully induced different levels of pressure on the operators, with the overall NASA-TLX score increasing with the difficulty of the N-back task (Figure 13b). This is consistent with the findings in the  $\alpha$  band's relative energy and the  $\theta/\alpha$  ratio, which also showed an increase in MWL with increasing N-back condition's difficulty.



(a) Median NASA-TLX scores for different CDIs

(b) Median NASA-TLX scores under different N-back tasks

**Figure 13.** Median NASA-TLX scores under different conditions.

#### 4.3. Limitations

The participants in this study were novice pilots, which could lead to easier mental overload than for experienced pilots under similar circumstances. To minimize the possibility of mental overload in situations requiring relatively low cognitive demands due to the participants' unskilled operations, we designed the tasks as cruise tasks and N-back tasks, which were easier to complete compared with other flight stages' tasks. The impact of the same CDI on pilots might also be different depending on their familiarity level with the cockpit. In future work, the participants can be pilots with various experience levels, the primary task can be a more cognitive resource-dependent one, such as takeoff, and the secondary tasks can be real parallel tasks, such as "communicating with an air traffic controller". This would be more similar to today's aviation and improve the study's universality.

The physiological evaluation tools for MWL could be more comprehensive in future studies. As reflected in the current results, HRV is not sensitive enough to detect changes in MWL when there are no significant changes. The eye-tracking technique can be considered a future evaluation tool. The pupil size change has been highlighted by previous studies as being sensitive enough for MWL measurement [65,66]. The participants' focused areas can be revealed by the areas of interest captured by the eye tracker [67,68], making it suitable for investigating the specific parts of the CDI that have a noticeable impact on participants' SA and MWL.

## 5. Conclusions

In this study, we compared the effect of using the G1000 Glass (GG) and Steam Gauge (SG) panels in the Cessna 172 series as cockpit display interfaces (CDIs). The results showed that the GG panel was more situational awareness (SA)-friendly, as the participants had lower mental workloads (MWLs) when using it. This emphasizes the importance of designing SA-friendly interfaces in the aviation industry. However, it is worth noting that MWL can also be affected by other parallel flight tasks involving a CDI.

This study also found that the N-back tasks used in the experiment were successful in inducing different levels of MWL in the participants. These findings have implications for the design of CDIs and the management of MWL in the aviation industry and contribute to all the areas that involve interface design and MWL-SA evaluation.

We believe that the results of this study provide a clear direction for future work, which can be replicated on a larger scale with more diverse groups to improve the representativeness of the study. Our next step is to investigate how the layout design affects MWL under different flight tasks such as takeoff and landing. Additionally, this study successfully demonstrated that measuring SA and MWL simultaneously is viable, providing a direction for future interface design and improvement by using these two aspects as evaluation tools together.

**Author Contributions:** Conceptualization, H.T. and M.P.; methodology, H.T. and M.P.; formal analysis, H.T.; data curation, H.T.; writing—original draft preparation, H.T.; writing—review and editing, B.G.L., D.T. and M.P.; supervision, M.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the UNNC Education Foundation through Li Dak Sum Innovation Fellowship under Grant LDS202307.

**Institutional Review Board Statement:** This study was conducted in accordance with the Declaration of Helsinki and approved by the UNNC Faculty of Science and Engineering's ethics committee.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The raw data supporting the conclusions of this article will be made available by the authors on request.

**Acknowledgments:** The authors acknowledge the support of the University of Nottingham Ningbo China (UNNC) Artificial Intelligence and Optimisation (AIOP) research group and the Sensors, Sensor Networks and Instrumentation (SSNI) research group. The authors also thank the School of Aerospace at the University of Nottingham Ningbo China, who provided access to the flight simulators used in this research.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Li, W.; Li, R.; Xie, X.; Chang, Y. Evaluating mental workload during multitasking in simulated flight. *Brain Behav.* **2022**, *12*, e2489. [[CrossRef](#)] [[PubMed](#)]
2. Wickens, C.D. Workload and situation awareness. In *Stress, Workload, and Fatigue*; Lawrence Erlbaum Associates, Inc.: Mahwah, NJ, USA, 2001; pp. 443–450.
3. Cain, B. *A Review of the Mental Workload Literature*; Technical Report; Defence Research And Development Toronto (Canada): North York, ON, Canada, 2007.

4. Wickens, C.D.; Hollands, J.G.; Banbury, S.; Parasuraman, R. *Engineering Psychology and Human Performance*; Psychology Press: London, UK, 2015; pp. 480–507.
5. Endsley, M.R. Toward a theory of situation awareness in dynamic systems. *Hum. Factors* **1995**, *37*, 32–64. [[CrossRef](#)]
6. Selcon, S.J.; Hardiman, T.D.; Croft, D.G.; Endsley, M.R. A Test-Battery Approach to Cognitive Engineering: To Meta-Measure or not to Meta-Measure, that is the Question! In Proceedings of the 40th Annual Meeting of Human Factors and Ergonomics Society, Santa Monica, CA, USA, 1 October 1996; SAGE Publications Sage CA: Los Angeles, CA, USA, 1996; Volume 40, pp. 228–232.
7. Stanton, N.A.; Plant, K.L.; Roberts, A.P.; Allison, C.K. Use of Highways in the Sky and a virtual pad for landing Head Up Display symbology to enable improved helicopter pilots situation awareness and workload in degraded visual conditions. *Ergonomics* **2019**, *62*, 255–267. [[CrossRef](#)] [[PubMed](#)]
8. Wei, H.; Zhuang, D.; Wanyan, X.; Wang, Q. An experimental analysis of situation awareness for cockpit display interface evaluation based on flight simulation. *Chin. J. Aeronaut.* **2013**, *26*, 884–889. [[CrossRef](#)]
9. Yan, S.; Wei, Y.; Tran, C.C. Evaluation and prediction mental workload in user interface of maritime operations using eye response. *Int. J. Ind. Ergon.* **2019**, *71*, 117–127. [[CrossRef](#)]
10. Feng, C.; Wanyan, X.; Liu, S.; Chen, H.; Zhuang, D. An approach to situation awareness (SA) assessment in flight simulation: SA dynamic circulation (SADC) model. *Hum. Factors Ergon. Manuf. Serv. Ind.* **2021**, *31*, 559–569. [[CrossRef](#)]
11. Casner, S.M. Perceived vs. measured effects of advanced cockpit systems on pilot workload and error: Are pilots' beliefs misaligned with reality? *Appl. Ergon.* **2009**, *40*, 448–456. [[CrossRef](#)] [[PubMed](#)]
12. Beringer, D.B.; Ball, J.D. A comparison of pilot navigation performance using conventional instrumentation, head-down, and head-up highway-in-the-sky primary flight displays. In Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting, Minneapolis, MN, USA, 8–12 October 2001; SAGE Publications Sage CA: Los Angeles, CA, USA, 2001; Volume 45, pp. 16–20.
13. Wickens, C.D. Situation awareness and workload in aviation. *Curr. Dir. Psychol. Sci.* **2002**, *11*, 128–133. [[CrossRef](#)]
14. Wickens, C.D. Multiple resources and performance prediction. *Theor. Issues Ergon. Sci.* **2002**, *3*, 159–177. [[CrossRef](#)]
15. Sharples, S.; Megaw, T. The definition and measurement of human workload. *Eval. Hum. Work.* **2015**, 516–544.
16. Andre, A.D.; Wickens, C.D.; Moorman, L. Display formatting techniques for improving situation awareness in the aircraft cockpit. *Int. J. Aviat. Psychol.* **1991**, *1*, 205–218. [[CrossRef](#)]
17. Wu, X.; Wanyan, X.; Zhuang, D.; Liu, S. Pilot situational awareness modeling for cockpit interface evaluation. In Proceedings of the International Conference on Engineering Psychology and Cognitive Ergonomics, Toronto, ON, Canada, 17–22 July 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 476–484.
18. Longo, L. Formalising Human Mental Workload as a Defeasible Computational Concept. Ph.D. Thesis, The University of Dublin, Trinity College, Ireland, 2014.
19. Morris, C.H.; Leung, Y.K. Pilot mental workload: How well do pilots really perform? *Ergonomics* **2006**, *49*, 1581–1596. [[CrossRef](#)] [[PubMed](#)]
20. Wei, Z.; Zhuang, D.; Wanyan, X.; Liu, C.; Zhuang, H. A model for discrimination and prediction of mental workload of aircraft cockpit display interface. *Chin. J. Aeronaut.* **2014**, *27*, 1070–1077. [[CrossRef](#)]
21. Wilson, G.F. An analysis of mental workload in pilots during flight using multiple psychophysiological measures. *Int. J. Aviat. Psychol.* **2002**, *12*, 3–18. [[CrossRef](#)]
22. Uenking, M.D.; Hughes, M.F. The Efficacy of using synthetic vision terrain-Textured images to improve pilot situation awareness. *SAE Trans.* **2002**, *111*, 572–583.
23. Davis, R.C.; Wilt, D.W.; Henion, J.; Alter, K.; Snow, P.; Deaton, J.E. Formal tests for LLM approaches using refined cockpit display technology. In *Enhanced and Synthetic Vision 2005*; SPIE: Bellingham, WA, USA, 2005; Volume 5802, pp. 231–238.
24. Lin, L.W.; Lu, M.S. Empirical research on the relationship between Helicopter Pilots' mental workloads and situation awareness levels. *J. Am. Helicopter Soc.* **2016**, *61*, 1–8. [[CrossRef](#)]
25. Wickens, C.D. Multiple resources and mental workload. *Hum. Factors* **2008**, *50*, 449–455. [[CrossRef](#)] [[PubMed](#)]
26. Horrey, W.J.; Wickens, J.D. Multiple resource modeling of task interference in vehicle control, hazard awareness and in-vehicle task performance. In Proceedings of the Driving Assessment Conference, Park City, UT, USA, 21–24 July 2003; Volume 2, pp. 7–12.
27. Wickens, C.D. Multiple resource time sharing models. In *Handbook of Human Factors and Ergonomics Methods*; CRC Press: Boca Raton, FL, USA, 2004; pp. 427–434.
28. Ryu, K.; Myung, R. Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *Int. J. Ind. Ergon.* **2005**, *35*, 991–1009. [[CrossRef](#)]
29. Liggett, K.K.; Reising, J.M.; Hartsock, D.C. Development and evaluation of a background attitude indicator. *Int. J. Aviat. Psychol.* **1999**, *9*, 49–71. [[CrossRef](#)]
30. Xiao, X.; Wanyan, X.; Zhuang, D. Mental workload prediction based on attentional resource allocation and information processing. *Bio-Med. Mater. Eng.* **2015**, *26*, S871–S879. [[CrossRef](#)]
31. Thorpe, A.; Estival, D.; Molesworth, B.; Eidl, A. Pilot errors: Communication comes last. *Saf. Sci.* **2022**, *149*, 105686. [[CrossRef](#)]
32. Rainieri, G.; Fraboni, F.; Russo, G.; Tul, M.; Pingitore, A.; Tessari, A.; Pietrantoni, L. Visual scanning techniques and mental workload of helicopter pilots during simulated flight. *Aerospace Med. Hum. Perform.* **2021**, *92*, 11–19. [[CrossRef](#)] [[PubMed](#)]
33. Endsley, M.R. Situation awareness. *Handb. Hum. Factors Ergon.* **2021**, *434*–455. [[CrossRef](#)]

34. Grether, W.F. Instrument reading. I. The design of long-scale indicators for speed and accuracy of quantitative readings. *J. Appl. Psychol.* **1949**, *33*, 363. [CrossRef]
35. Jaeggi, S.M.; Buschkuhl, M.; Perrig, W.J.; Meier, B. The concurrent validity of the N-back task as a working memory measure. *Memory* **2010**, *18*, 394–412. [CrossRef] [PubMed]
36. Anderson, V.L.; McLean, R.A. *Design of Experiments: A Realistic Approach*; CRC Press: Boca Raton, FL, USA, 2018; Chapter 8.
37. Lean, Y.; Shan, F. Brief review on physiological and biochemical evaluations of human mental workload. *Hum. Factors Ergon. Manuf. Serv. Ind.* **2012**, *22*, 177–187. [CrossRef]
38. Hankins, T.C.; Wilson, G.F. A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight. *Aviat. Space Environ. Med.* **1998**, *69*, 360–367. [PubMed]
39. Fink, A.; Grabner, R.; Neuper, C.; Neubauer, A. EEG alpha band dissociation with increasing task demands. *Cogn. Brain Res.* **2005**, *24*, 252–259. [CrossRef] [PubMed]
40. Brouwer, A.M.; Hogervorst, M.A.; Van Erp, J.B.; Heffelaar, T.; Zimmerman, P.H.; Oostenveld, R. Estimating workload using EEG spectral power and ERPs in the n-back task. *J. Neural Eng.* **2012**, *9*, 045008. [CrossRef]
41. Esposito, F.; Aragri, A.; Piccoli, T.; Tedeschi, G.; Goebel, R.; Di Salle, F. Distributed analysis of simultaneous EEG-fMRI time-series: Modeling and interpretation issues. *Magn. Reson. Imaging* **2009**, *27*, 1120–1130. [CrossRef]
42. Jasper, H.H. The ten-twenty electrode system of the International Federation. *Electroencephalogr. Clin. Neurophysiol.* **1958**, *10*, 370–375.
43. Jahn, N.T.; Meshi, D.; Bente, G.; Schmälzle, R. Media neuroscience on a shoestring: Examining electrocortical responses to visual stimuli via mobile EEG. *J. Media Psychol. Theor. Methods Appl.* **2022**. [CrossRef]
44. Pu, L.; Lion, K.M.; Todorovic, M.; Moyle, W. Portable EEG monitoring for older adults with dementia and chronic pain—A feasibility study. *Geriatr. Nurs.* **2021**, *42*, 124–128. [CrossRef] [PubMed]
45. Jaiswal, D.; Chowdhury, A.; Banerjee, T.; Chatterjee, D. Effect of Mental Workload on Breathing Pattern and Heart Rate for a Working Memory Task: A Pilot Study. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 2202–2206.
46. Murai, K.; Hayashi, Y.; Okazaki, T.; Stone, L.C.; Mitomo, N. Evaluation of ship navigator’s mental workload using nasal temperature and heart rate variability. In Proceedings of the 2008 IEEE International Conference on Systems, Man and Cybernetics, Singapore, 12–15 October 2008; IEEE: Piscataway, NJ, USA, 2008; pp. 1528–1533.
47. Murata, A. Experimental discussion on measurement of mental workload—evaluation of mental workload by HRV measures. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* **1994**, *77*, 409–416.
48. Chattopadhyay, S.; Das, R. Comparing heart rate variability with polar H10 sensor and pulse rate variability with LYFAS: A novel study. *J. Biomed. Eng. Technol.* **2021**, *9*, 1–9. [CrossRef]
49. H-Aviation. Cessna 172S Skyhawk—FAQ’s. 2020. Available online: <https://www.h-aviation.com/cessna-172s---skyhawk.html#:~:text=How%20fast%20does%20the%20Cessna,a%20maximum%20of%20302km%2Fh> (accessed on 10 November 2020).
50. Hart, S.G. NASA-task load index (NASA-TLX); 20 years later. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, San Francisco, CA, USA, 16–20 October 2006; Sage Publications Sage CA: Los Angeles, CA, USA, 2006; Volume 50, pp. 904–908.
51. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology*; Elsevier: Amsterdam, The Netherlands, 1988; Volume 52, pp. 139–183.
52. Shankar, S.; Singh, R. Demystifying statistics: How to choose a statistical test? *Indian J. Rheumatol.* **2014**, *9*, 77–81. [CrossRef]
53. Khorana, A.; Pareek, A.; Ollivier, M.; Madjarova, S.J.; Kunze, K.N.; Nwachukwu, B.U.; Karlsson, J.; Marigi, E.M.; Williams III, R.J. Choosing the appropriate measure of central tendency: Mean, median, or mode? *Knee Surgery Sport. Traumatol. Arthrosc.* **2023**, *31*, 12–15. [CrossRef]
54. Wickens, C.D. Attentional tunneling and task management. In Proceedings of the 2005 International Symposium on Aviation Psychology, Dayton, OH, USA, 18–21 April 2005; p. 812.
55. Champseix, R. Hrv-Analysis, 2021. Version 1.0.4. Github. 2021. Available online: <https://github.com/Aura-healthcare/hrv-analysis> (accessed on 4 November 2020).
56. Veltman, J.; Gaillard, A. Physiological workload reactions to increasing levels of task difficulty. *Ergonomics* **1998**, *41*, 656–669. [CrossRef] [PubMed]
57. Yao, D.; Wang, L.; Oostenveld, R.; Nielsen, K.D.; Arendt-Nielsen, L.; Chen, A.C. A comparative study of different references for EEG spectral mapping: The issue of the neutral reference and the use of the infinity reference. *Physiol. Meas.* **2005**, *26*, 173. [CrossRef]
58. Mallat, S.G. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11*, 674–693. [CrossRef]
59. Hazarika, N.; Chen, J.Z.; Tsoi, A.C.; Sergejew, A. Classification of EEG signals using the wavelet transform. *Signal Process.* **1997**, *59*, 61–72. [CrossRef]
60. Tsang, P.; Psychologist, M. *Mental Workload and Situation Awareness*; Wiley: Hoboken, NJ, USA, 2006; Chapter 9, pp. 243–268.
61. Napierala, M.A. What is the Bonferroni correction? *Aaos Now* **2012**, *40*–41. Available online: <http://www.aaos.org/news/aaosnow/apr12/research7.asp> (accessed on 6 June 2022).

62. Kandel, E.R.; Schwartz, J.H.; Jessell, T.M.; Siegelbaum, S.; Hudspeth, A.J.; Mack, S. *Principles of Neural Science*, 6th ed.; McGraw-Hill: New York, NY, USA, 2021; p. 1450.
63. Pritschet, L.; Powell, D.; Horne, Z. Marginally significant effects as evidence for hypotheses: Changing attitudes over four decades. *Psychol. Sci.* **2016**, *27*, 1036–1042. [[CrossRef](#)] [[PubMed](#)]
64. Stipacek, A.; Grabner, R.; Neuper, C.; Fink, A.; Neubauer, A. Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load. *Neurosci. Lett.* **2003**, *353*, 193–196. [[CrossRef](#)] [[PubMed](#)]
65. Iqbal, S.T.; Zheng, X.S.; Bailey, B.P. Task-evoked pupillary response to mental workload in human-computer interaction. In Proceedings of the CHI’04 Extended Abstracts on Human Factors in Computing Systems, Vienna, Austria, 24–29 April 2004; pp. 1477–1480.
66. Marquart, G.; Cabrall, C.; De Winter, J. Review of eye-related measures of drivers’ mental workload. *Procedia Manuf.* **2015**, *3*, 2854–2861. [[CrossRef](#)]
67. Jacob, R.J. Eye tracking in advanced interface design. *Virtual Environ. Adv. Interface Des.* **1995**, *258*, 2.
68. Poole, A.; Ball, L.J. Eye tracking in HCI and usability research. In *Encyclopedia of Human Computer Interaction*; IGI Global: Hershey, PA, USA, 2006; pp. 211–219.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.