The Aeronautical Journal (2025), **129**, pp. 609–626
doi:10.1017/aer.2024.103

RESEARCH ARTICLE

Visual cognition-based optimised design of primary flight displays in cockpits

Z. Zeng[1,2], Y. Sun[1,2], X. Liu[1,3], Y. Jie[1,3] and Y. Zeng[1,2]

[1]College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing, China
[2]State Key Laboratory of Airliner Integration Technology and Flight Simulation, Shanghai, China
[3]Flight Performance Division, Shanghai Aircraft Airworthiness Certification Center, Shanghai, China
Corresponding author: Y. Sun; Email: sunyc@nuaa.edu.cn

Received: 30 April 2024; **Revised:** 4 September 2024; **Accepted:** 5 September 2024

Keywords: cockpit primary flight display interface; design factors; interface optimisation; visual cognition

Abstract

The objective of this study was to investigate the effects of different cockpit primary flight display (PFD) interface designs on pilot cognitive efficiency and cognitive load. This study designed five optimised PFD interfaces and conducted interface cognition experiments to assess cognitive responses across six different PFD interface designs, including the original design. It compared various subjective and objective metrics across different interface designs and evaluated the impact of each design factor on cognitive task performance. The experimental results show that the PFD interface in the original interface design performs better under different flight symbol designs, and the interface with 50% increase in font size performs better among interface designs with different font sizes with relatively lower cognitive load. This study provides experimental support and optimization suggestions for the optimal design of cockpit PFD interface, which can help improve pilots' perception and operational capabilities, and thus enhance task performance efficiency and flight safety. Future research can investigate the effects of various design factors on the cognitive effects of the interface to enhance the ongoing improvement and optimisation of interface design.

Nomenclature

ANN	artificial neural network
bps	blinks per second
EEG	electroencephalography
HWD	head-worn display
NASA-TLX	NASA task load index
PFD	primary flight display
PSD	power spectral density
RLD	relative layout design
SUS	system usability scale
SVM	support vector machine
SWAT	subjective workload assessment technique

1.0 Introduction

Pilots need to continuously interact with the cockpit display interface while performing flight tasks [1–3]. This interaction enables them to access information about the aircraft's flight status and surrounding situations, facilitating effective aircraft manoeuvering [4, 5]. The primary flight display (PFD) is a crucial component of cockpit human-computer interaction. Its primary function is to transmit flight information and display the aircraft's flight status in real time [6]. The PFD interface provides pilots with the most critical information for safe flight. The main elements of the interface include the airspeed indicator, altitude indicator, attitude indicator, heading indicator and other similar instruments, which provide pilots with real-time flight information such as airspeed, altitude, pitch and roll angle and heading [7, 8]. The PFD interface enables the pilot to obtain accurate information and to perform the requisite operations in an efficient manner, thereby facilitating the completion of the flight task [9, 10, 19].

Pilots' work efficiency and performance are significantly influenced by a multitude of factors, such as their psychological and physiological states [11], the extent and quality of their training and experience [12], environmental conditions [13], the design and layout of the cockpit display interfaces [19], as well as the complexity of the tasks. Among these factors, the cockpit display interface design has a direct impact on pilots' information acquisition and processing, which ultimately affects their driving operations [14–16]. Therefore, the design and layout of the PFD interface are crucial to the pilot's efficiency and understanding of the flight status. It is one of the important factors that affect the pilot's efficient execution of the task. Good interface design can help the pilot quickly and accurately obtain key information, improving the accuracy and responsiveness of the flight decision [17–19].

The optimal design of cockpit display interfaces has become an important scientific issue due to the rapid development of aviation technology and the increasing complexity of flight missions. This is crucial to ensure pilots' safe flight and mission performance optimisation. Li et al. [20] utilised an improved visualisation design to optimise the cockpit PFD interface. This was achieved by highlighting parameters through green borders or associating them with relevant flight mode changes. Şenol [21] proposed the relative layout design (RLD) optimisation model for conventional cockpit interfaces. The RLD model optimises the position of indicators on the interface, improving display interface usability and contributing to pilot-aircraft interaction and flight safety. Zhang et al. [22] conducted an experimental study on the scale intervals, horizontal and vertical distances of different scale bands of the head-worn display (HWD). They obtained an optimised layout of the scale bands of the HMD interface with better task performance.

The optimisation of cockpit display interfaces can be achieved by applying the principles of human-computer interaction. The principles of human-computer interaction primarily include usability, user experience, information presentation and feedback [23–26]. These principles aim to enhance system ease of use and user satisfaction. There also are a few studies conducted in the past on the optimisation and design of human-computer interaction interface factors. Shen et al. [27] conducted a study on the impact of color combination, luminance contrast and icon area ratio on the visual search of graphical symbols. The study evaluated the significance of each factor on the visual search of graphical symbols by measuring the accuracy and response time of the icon search task. Rettenmaier et al. [28] conducted two experiments to investigate the readability of different content types (text and symbols) and colours at varying distances. The results showed that, at a fixed distance, text must be larger than symbols to maintain readability. Yang et al. [29] introduced the artificial neural network (ANN) and support vector machine (SVM) algorithm models into the study of the optimal design of automotive T-panels. The models were used to identify design features of automotive panels and detect system usability, establishing a relationship between the design features of the vehicle panels and the system usability. Dou et al. [30] proposed an extended analysis and optimisation method for interface elements to solve the contradiction between content colour and driver visual fatigue in augmented reality heads-up display interfaces. They constructed a content colour selection model for two conditions and described the distance from the selected colour to the optimal interval to better balance the driver's cognitive load and situational awareness.

The key to the design of human-machine interfaces lies in the achievement of seamless interaction, which is aimed at reducing the cognitive load on the user and enhancing their perception and operational abilities [31–33]. Therefore, when designing PFD interfaces, it is necessary to consider enhancing pilots' comfort during flight and reducing cognitive load. The specific design process can be informed by the principles of human-machine interface design, which mainly focus on the visual information such as interface layout, colour, graphics and text [34–37]. The interface layout should be logical and straightforward, the use of colour should optimize visual impact and the manner of conveying information in the interface, such as symbols, text and colour, should be intuitive, accessible and conducive to learning.

However, although some progress has been made in previous research, the complexity of the factors affecting pilot cognitive ergonomics and the lack of clarity on how interface design factors affect the visual cognitive effects of pilots have posed a challenge to the optimisation of PFD interface design. In this study, two design factors of the cockpit PFD interface are optimised. The aircraft symbol optimisation aims to reflect the aircraft attitude more clearly and intuitively to improve the pilot's cognitive efficiency and reduce cognitive load. The font size of the displayed information is optimised to improve the readability and legibility of the flight data, thus reducing the pilot's cognitive load when accessing critical information.

In this study, an experimental evaluation method was used to verify the ergonomics of the optimised PFD interface and compare it to the original interface before optimisation. An experiment was conducted to obtain participants' subjective feedback, task performance, and physiological indicators to evaluate the ergonomics of the PFD interface before and after optimisation and the pilot's workload, and to determine whether the optimised interface has the potential to improve pilot efficiency and reduce erroneous decisions, and whether it has a significant improvement in pilots' perception and understanding.

2.0 Method

2.1 Participants

Based on the logic of participant selection in Ref. [38], 11 male flight cadets were recruited as participants for the study. Their average age was 24.3 years (SD = 3.0), and they all had extensive experience with simulated cockpit flights. They were required to be in good health, with no history of neurological disorders or genetic conditions, and have normal colour vision (without colour blindness or deficiency) and visual acuity (or corrected visual acuity) of 1.0 or above. In addition, all participants were very familiar with the composition of the cockpit PFD interface and were proficient in using it to determine flight status and various flight parameters.

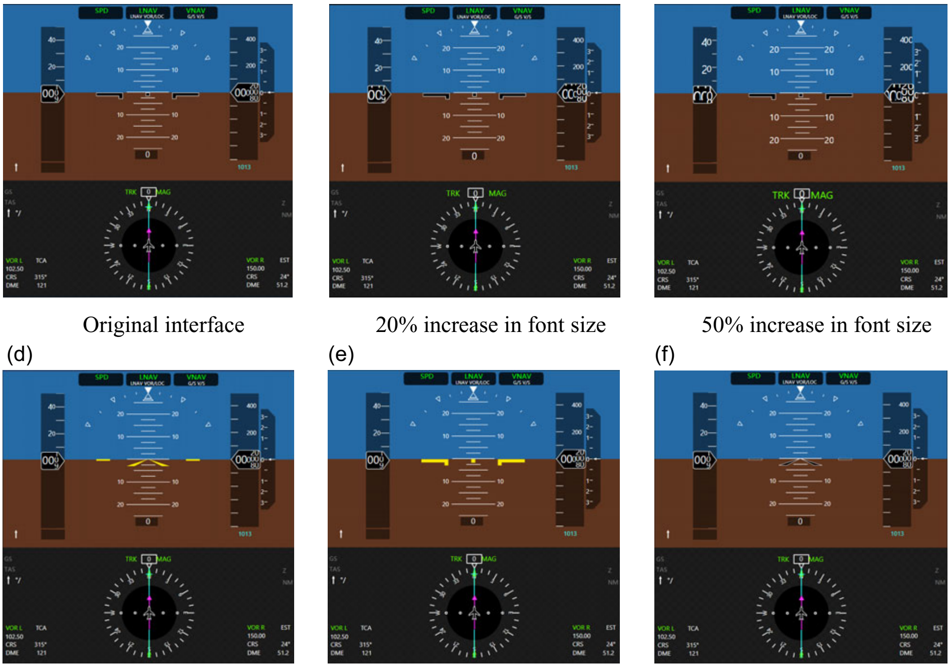
2.2 Materials

In this study, all the designed PFD interfaces were employed as experimental material. There were six designs, including the original, which primarily varied in graphical symbols and font sizes of the displayed information. These five optimised interfaces were designed based on airworthiness regulations and human-machine interface design principles [11, 12]. Modifications to the shapes and colours symbol, and increased font sizes aimed to more clearly and intuitively reflect the aircraft's flight status and flight instrument data. Figure 1 displays all the stimulus interface materials used in the test.

2.3 Apparatus and environment

PsychoPy software was used to control the trial flow and present interface stimuli in this experiment. PsychoPy is a widely used open-source software for experimental design and data collection in psychology and neuroscience. It enables the full flow of this experiment to be designed, controlling the random appearance of all dynamic stimulus interfaces and cognitive questions. Furthermore, it allows the recording of the participants' reaction time and accuracy for each interface cognition.

The SMI ETG eye-tracker and ANT Neuro electroencephalograph (EEG) were used to record the eye-movement and electroencephalographic activities of the participants during the experimental task, which allowed for analysis of the cognitive processes and allocation of attention during the trial. The physical drawings of the SMI ETG eye-tracker and ANT Neuro EEG used in the experiment are shown in Fig. 2.



(a) Original interface (b) 20% increase in font size (c) 50% increase in font size
(d) Symbol 1 (e) Symbol 2 (f) Symbol 3

Figure 1. *Experimental Material: PFD Interface.*

(a) The SMI ETG eye-tracker (b) The ANT Neuro EEG

Figure 2. *The SMI ETG eye-tracker and ANT Neuro EEG, photographed by the authors.*

To ensure the stability of the experimental environment and the comfort of the participants, this experiment was set up in a quiet, spacious laboratory. Participants were seated in a comfortable chair facing a computer screen for the interface cognition experiment. The screen used was a 14-inch LCD monitor with a resolution of 1920x1080 pixels and the viewing distance was 60 cm. In addition, the lighting intensity in the laboratory was controlled to ensure adequate brightness and uniformity. The average brightness of the screen and the room were maintained at approximately 500 lux and 300 lux, respectively, to provide a consistent visual experience for the participants.

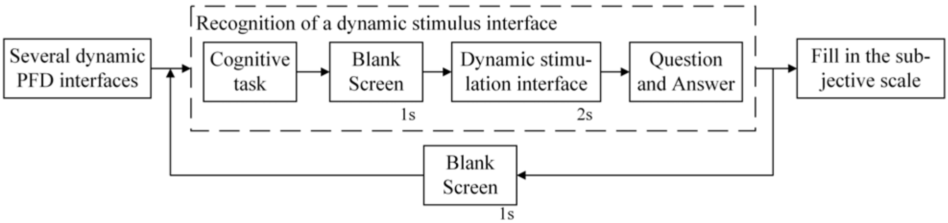


Figure 3. *Cognitive task trials for a single PFD interface design: A description of the flow.*

2.4 Task and procedure

This experiment is an interface evaluation test based on a multi-factorial design with the aim of evaluating the cockpit PFD interface and selecting the optimal interface design from five optimised designs proposed in Section 2.2. The independent variables are different cockpit PFD interface designs and different cognitive tasks for displaying the interface, while the dependent variables are the performance, psychological feelings and physiological data of the participants when performing the tasks.

The experimental task required the participants to observe the cockpit PFD interface displayed on the computer and to complete a cognitive task for each stimulus interface to judge the flight status of the aircraft and each flight parameter. To better simulate the real flight and to make the test results reliable and valid, the cockpit PFD interfaces presented to the participants were dynamic stimulus interfaces with changes in flight instruments. The cognitive flow of a PFD interface is shown in Fig. 3.

In this experiment, several dynamic stimulus interfaces were prepared as experimental materials for the cognitive task and these stimulus materials were presented randomly to the participants during the experiment. Prior to the presentation of the dynamic stimulus interface, participants were required to determine the cognitive task to be completed. All cognitive tasks set for this experiment are shown in Table 1. For each dynamic stimulus interface, participants were required to randomly complete one cognitive task. Following a one-second interval, the interface stimulus was then displayed. Participants were able to recognize the dynamic stimulus interface in accordance with the cognitive task corresponding to the test questions and answer options. They were then required to press the A, B, C or D key to answer the questions. To ensure consistency among all participants, each cognitive task was paired with a set of fixed options. The correct answers to these questions were predetermined based on the specific content of different dynamic PFD interface materials and the questions themselves. Following the response to the question, there was a one-second blank screen before participants proceeded to the subsequent dynamic stimulus interface for recognition.

Once the answer had been entered, a one-second blank screen was presented before the next dynamic stimulus interface of the recognition commenced. Following the response to the question, a one-second interval was observed before participants proceeded to the subsequent dynamic stimulus interface for recognition. After completing the cognitive tasks for all dynamic stimulus interfaces within the current interface design, participants were required to complete two subjective scales: the NASA-TLX and the SUS. These scales were used to subjectively evaluate the current interface design in terms of workload and usability, respectively (see Section 2.5.2 for a detailed description). Finally, a three-minute break was taken and then the cognitive test of the next interface design was carried out until the cognition of all the designed interfaces had been completed.

When performing flight tasks, pilots need to constantly monitor the flight attitude, airspeed, altitude, vertical speed and other key information, to help them maintain proper flight state, ensure flight safety and avoid risks [39–41]. A total of six interface cognitive tasks were devised based on the key display information of the PFD interface in the experiment. These tasks were accompanied by corresponding test questions, as shown in Table 1. Furthermore, changes in interface symbols primarily affect a pilot's cognition of three types of flight information: flight state, pitch angle and roll direction, while changes in interface font sizes primarily affect a pilot's cognition of flight state, airspeed, altitude and heading.

Table 1. *Interface cognition tasks and testing issues*

Number	Task	Questions	Options
1	Flight state: Determine which phase of flight the aircraft is in	Determine the flight condition corresponding to the current interface.	A. Climbing B. Descending C. Flying level D. Turning
2	Airspeed: Determine which airspeed value is closest	The following is closest to the speed value of the current interface.	A. 150 B. 300 C. 450 D. 600
3	Altitude: Determine which altitude value is closest	The following is closest to the altitude value of the current interface.	A. 2000 B. 4000 C. 6000 D. 8000
4	Heading: Determine which heading angle is closest	The following is closest to the heading angle of the current interface.	A. 0/360 B. 90 C. 180 D. 270
5	Pitch angle: Determine the pitch state of the aircraft	Determine the corresponding pitch setting of the current interface.	A. Head up B. Head down C. Flying level
6	Roll direction: Determine if the aircraft has rolled and the direction of the roll	Determine the roll direction according to the current interface.	A. Roll to the left B. Roll to the right C. No roll

Therefore, when designing the cognitive tasks corresponding to the different stimulus interfaces, the participants were asked to recognise the flight state, pitch angle and roll direction of the interface under the different interface symbols and to recognise the flight state, airspeed, altitude and heading of the interface under the different interface font sizes.

2.5 Data recording and analysis

2.5.1 Task performance

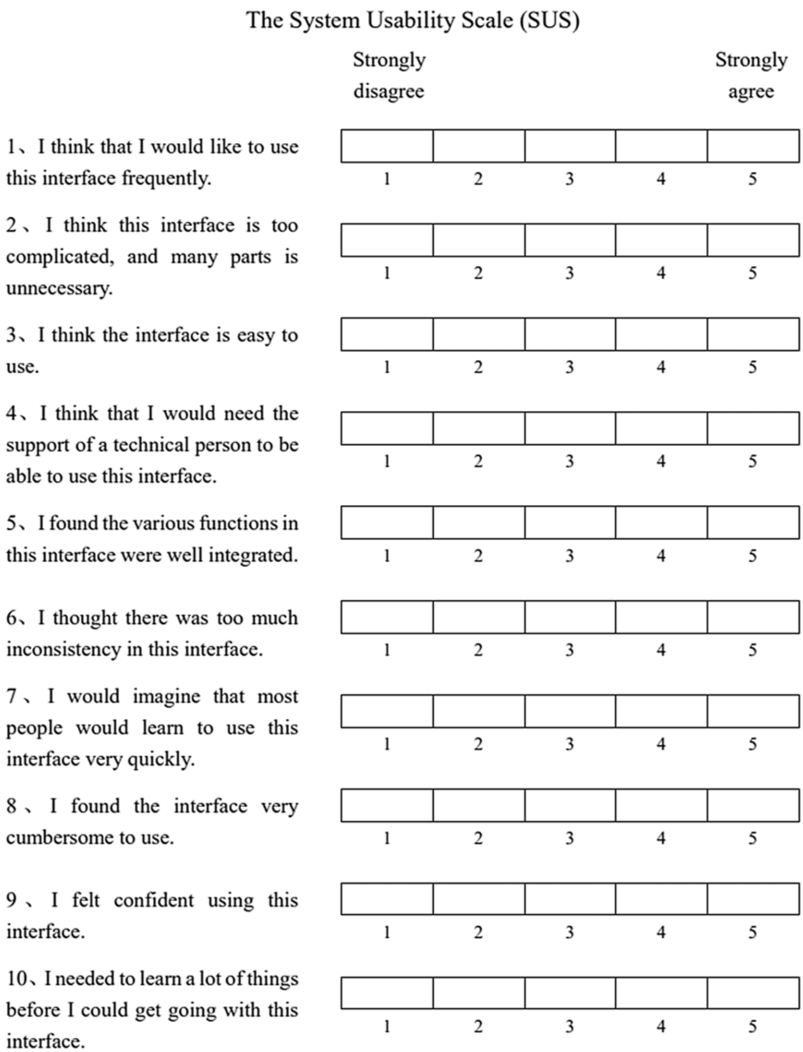
Task performance assessment is an objective yet direct measure of cognitive load that evaluates a subject's level of cognitive load through the quality of their task completion [42]. In general, the lower the cognitive load of a task, the higher the level of task performance and the less time spent [43].

In this experiment, reaction time and accuracy are used as performance evaluation indices to assess the ergonomics of the interface. Participants with faster reaction times and higher accuracy in recognising the interface have a lower cognitive load, indicating better ergonomic design.

2.5.2 Subjective evaluation

The assessment of the interface requires participants to recall their subjective experiences during the experimental task and evaluate the level of cognitive load in terms of mental effort, task difficulty and time pressure. The cognitive load level of the task is then assessed in terms of mental effort, task difficulty and time pressure. The most commonly used subjective assessment scales at present include the NASA Task Load Index (NASA-TLX), the System Usability Scale (SUS), the Likert scale, the Subjective Workload Assessment Technique (SWAT) scale, and others [44–46]. The experiment evaluated the optimisation effect of the cockpit PFD interface using SUS and NASA-TLX scale as subjective assessment tools. Participants completed the scales based on their subjective feelings after each cognitive task of the interface design.

The NASA-TLX scale is a widely used subjective workload assessment tool designed to evaluate perceived workload in human-machine interaction tasks. As a multi-dimensional questionnaire, this



The System Usability Scale (SUS)

	Strongly disagree				Strongly agree
1、I think that I would like to use this interface frequently.	1	2	3	4	5
2、I think this interface is too complicated, and many parts is unnecessary.	1	2	3	4	5
3、I think the interface is easy to use.	1	2	3	4	5
4、I think that I would need the support of a technical person to be able to use this interface.	1	2	3	4	5
5、I found the various functions in this interface were well integrated.	1	2	3	4	5
6、I thought there was too much inconsistency in this interface.	1	2	3	4	5
7、I would imagine that most people would learn to use this interface very quickly.	1	2	3	4	5
8、I found the interface very cumbersome to use.	1	2	3	4	5
9、I felt confident using this interface.	1	2	3	4	5
10、I needed to learn a lot of things before I could get going with this interface.	1	2	3	4	5

Figure 4. *The System Usability Scale (SUS).*

scale evaluates subjective workload from six perspectives: mental demand, physical demand, temporal demand, performance, effort and frustration [47]. In this experiment, the NASA-TLX scale completed by participants was used to evaluate their workload during cognitive tasks for each interface design. The total score of the scale was obtained by summing the six individual ratings, each ranging from 0 to 20, with higher scores indicating higher subjective workload.

The SUS is commonly used to assess the usability of various products and systems. It consists of 10 items, each rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). In this experiment, the SUS scale used was an improved version of the original SUS [48] to better suit the specific requirements of the study and to more accurately assess the usability of each interface design (see Fig. 4). Additionally, the SUS has a predefined procedure for calculating the total SUS score as follows [49]: First, the scores of the 10 items are transformed. For odd-numbered items, subtract 1 from the score; for even-numbered items, subtract the score from 5. Then, sum all the transformed scores to get a total score. Finally, multiply this total score by 2.5 to obtain the overall SUS score. This calculation process can be expressed by the following formula:

$$SUS_{score} = 2.5\left(\sum_{i\in\{1,3,5,7,9\}}(x_i - 1) + \sum_{i\in\{2,4,6,8,10\}}(5 - x_i)\right), \tag{1}$$

where x_i represents the score for the i-th item. After this transformation, the total SUS score ranges from 0 to 100, with higher scores indicating better system usability.

2.5.3 Objective evaluation

Objective assessment is used to evaluate the cognitive load level through the changes of human physiological indices [46, 50]. Previous studies have demonstrated that physiological indicators such as electroencephalography (EEG) and eye-tracking indicators vary with cognitive load levels, and by collecting and analyzing data on these indices, cognitive load levels can be inferred [51, 52]. The experiment collected EEG and eye movement physiological indicators to assess the cognitive load of subjects during the interface recognition test task. The optimal design for each interface was selected by comparing the measurement results.

EEG signals can be divided into five frequency bands, namely alpha, beta, theta, delta and gamma bands [53, 54]. When EEG frequency domain characteristics are analyzed, the activities of alpha and theta bands are usually selected to analyze the brain activity patterns of subjects [55]. In this experiment, the parietal alpha band power spectral density (PSD) and the frontal theta band power spectral density in the frequency domain characteristics of the EEG signal were selected as evaluation indices, and in general, the parietal alpha band PSD decreases with the increase of brain load, while the frontal theta band PSD increases with the increase of brain load. By analyzing the parietal alpha band PSD and the frontal theta band PSD, we can compare the brain load of participants exposed to different interface designs.

Regarding the eye movement indicators, some scholars have found that the indicators related to the pilot's cognitive load mainly include: pupil diameter, number of gaze points, gaze duration, blink rate [56–58]. Since the random cognitive tasks are explained in advance before cognizing the interface in the experiment, and the subjects already know which instrument they need to cognize in the interface, the eye movement indicators such as gaze duration and number of gaze points are not of much reference value to the results, and they cannot effectively differentiate the subjects' cognitive load under different interfaces and cognitive tasks, so we choose blink rate and pupil diameter to analyze the subjects' cognitive load, and then we choose the physiological eye movement indicators. Therefore, blink rate and pupil diameter were chosen to analyze the cognitive load of subjects under different interfaces and cognitive tasks. In general, the blink rate is inversely proportional to the workload, and the blink rate decreases with the increase of physiological load such as visual load, and the increase of load may also cause physiological and emotional responses such as arousal and excitement, which leads to the increase of pupil diameter [59, 60].

3.0 Result

3.1 Task performance

After excluding invalid data with extreme outliers in reaction time and accuracy, the average values of reaction time and accuracy for all participants on different cognitive tasks in interface with different symbols and font sizes are presented in Fig. 5.

From the results, it could be seen that the original interface had the best task performance in the cognitive task of flight state with the least reaction time and the highest accuracy rate, while the interface with symbol 1 had the best task performance in the cognitive task of roll direction. In addition, the interface with symbol 1 had higher accuracy than the other three interfaces in the cognitive task of pitch angle, but the interface with symbol 3 had the least reaction time among the four interfaces (Table 3).

The cognitive accuracy of each task was basically at a high level across all three interface designs with different font size. The original interface had the highest accuracy in airspeed and altitude cognition

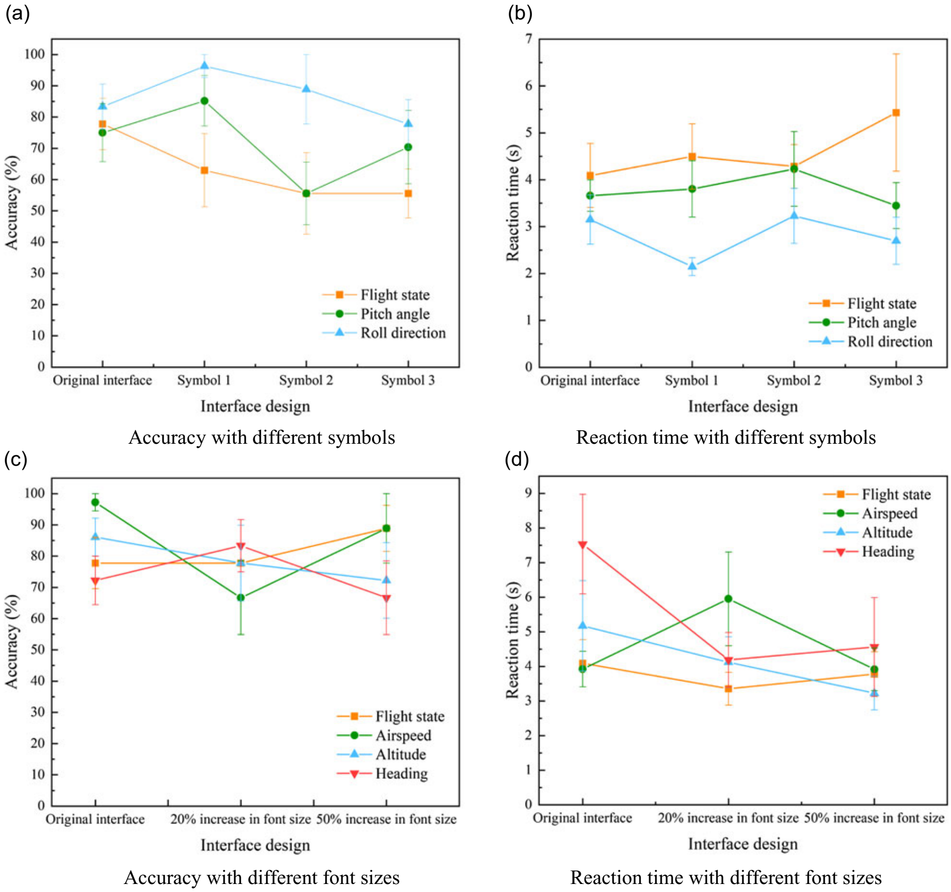


Figure 5. *The differences in reaction time and accuracy across cognitive tasks among six interface designs with different aircraft symbols and font sizes.*

task, and the interface with a 20% font enlargement had the highest accuracy in cognition of heading angle, while the interface with a 50% font enlargement had the highest accuracy in flight state cognition. Apart from the airspeed cognitive task with a 20% font enlargement, the reaction time of the participants was shorter when using the enlarged interface compared to the original. Furthermore, most cognitive tasks had a shorter reaction time with a 50% font enlargement of the interface (Table 4).

3.2 Subjective evaluation

For the NASA-TLX and SUS scales, Table 2 shows the mean values of the total scores for all valid participants under different interface designs. To highlight the optimal interface design, we used the following annotation method: bold values in the NASA-TLX scores indicate the lowest score, representing the design with the least cognitive load; bold values in the SUS scores indicate the highest score, representing the design with the highest user satisfaction.

According to the table, the interface with symbol 3 had the lowest total scores on the NASA-TLX and highest scores on the SUS among the PFD interface with different aircraft symbols. Among the interface designs with different font sizes, the table shows that font enlargement led to lower total scores on the NASA-TLX and the SUS compared to the original interface. The interface with a 20% font enlargement

Table 2. *Scale score results*

Scale	NASA-TLX	SUS
Original interface	55.8±15.30	**66.56±11.41**
Symbol 1	57.5±15.55	63.06±15.71
Symbol 2	49.4±13.35	61.94±15.27
Symbol 3	**49.3±17.92**	**67.22±12.66**
20% increase in font size	**54±12.22**	66.11±15.64
50% increase in font size	55.2±19.02	63.78±17.14

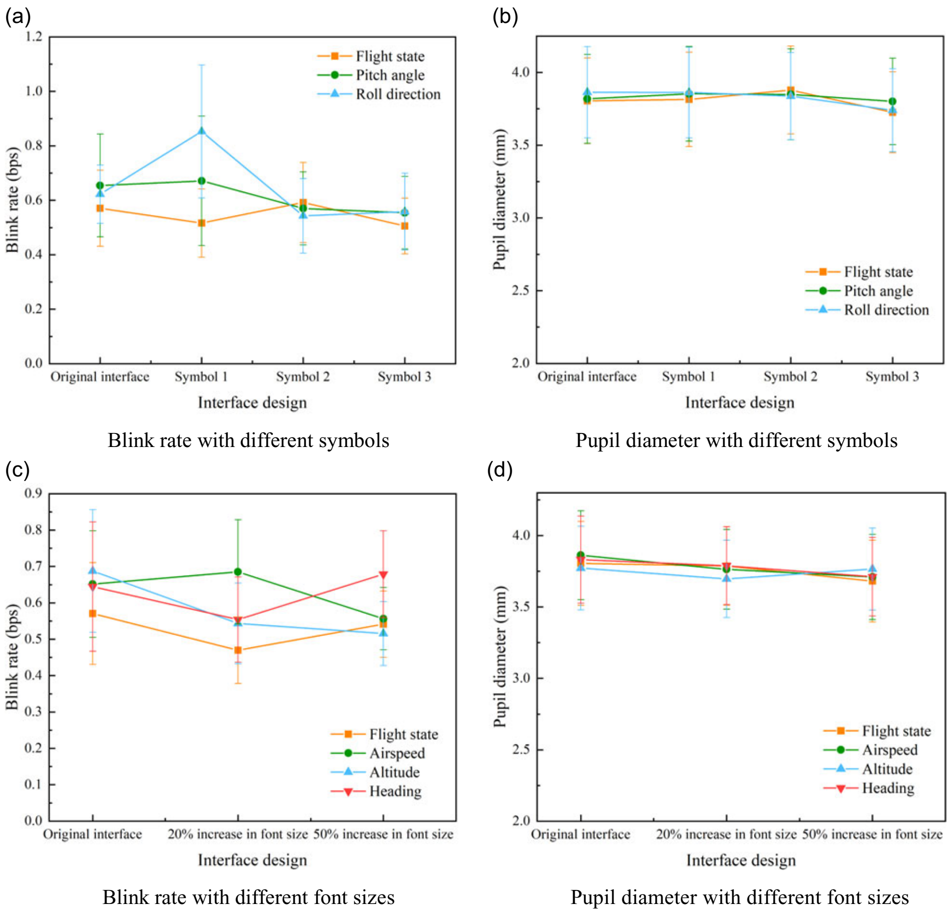


Figure 6. *The differences in blink rate and pupil diameter across cognitive tasks among six PFD interface with different aircraft symbols and font sizes.*

received the lowest NASA-TLX total score, while the original interface design received the highest SUS scale score.

3.3 Eye movement features

Figure 6 shows the average values of blink rate and pupil diameter for all participants after excluding invalid data in different interface designs and different cognitive tasks.

Table 3. *Mean values of effective feature metrics were calculated for the interface with different aircraft symbols and cognitive tasks*

		Flight state	Pitch angle	Roll direction
Accuracy (%)	Original interface	**77.78**	75.00	83.33
	Symbol 1	62.96	**85.19**	**96.30**
	Symbol 2	55.56	55.56	88.89
	Symbol 3	55.56	70.37	77.78
Reaction time (s)	Original interface	**4.0912**	3.6608	3.1509
	Symbol 1	4.4955	3.8035	**2.1468**
	Symbol 2	4.2834	4.2291	3.2283
	Symbol 3	5.4319	**3.4468**	2.6969
Parietal Alpha PSD ($\times 10^{-10}$ μV^2/Hz)	Original interface	3.9987	4.0605	**4.6995**
	Symbol 1	3.9263	3.5692	3.8047
	Symbol 2	4.2066	3.9311	4.3602
	Symbol 3	**4.6017**	**4.2903**	4.4422
Frontal Theta PSD ($\times 10^{-8}$ μV^2/Hz)	Original interface	**4.5312**	5.4931	**3.9337**
	Symbol 1	5.1672	**4.1912**	5.4667
	Symbol 2	4.6034	4.3411	4.3691
	Symbol 3	5.1723	4.8543	5.0918
Blink rate (bps)	Original interface	0.5707	0.6543	0.6224
	Symbol 1	0.5164	**0.6714**	**0.8527**
	Symbol 2	**0.5917**	0.5700	0.5429
	Symbol 3	0.5055	0.5543	0.5582
Pupil diameter (mm)	Original interface	3.8052	3.8181	3.8639
	Symbol 1	3.8148	3.8532	3.8622
	Symbol 2	3.8790	3.8492	3.8367
	Symbol 3	**3.7260**	**3.8010**	**3.7403**

*bold text indicates that these evaluation metrics performed relatively better when participants read the corresponding interface.

Based on the result, it was found that when recognizing the flight state corresponding to the PFD interface, the average blink rate and average pupil diameter of participants were the largest in the interface with symbol 2 and the smallest in the interface with symbol 3. When participants recognized the pitch angle of the interface, their average blink rate and average pupil diameter was the with symbol 1, indicating that the participants had a lower cognitive load in the interface design with symbol 1. This was evidenced by the fact that the average blink rate and average pupil diameter of the original interface were the largest, indicating that the participants had a lower cognitive load under the interface design with symbol 1. The rolling direction of the interface with symbol 1 was most easily recognized, with the highest average blink rate and average pupil diameter. However, the average pupil diameter with symbol 1 did not differ significantly from the original interface, indicating that the participants had an easier time and a lower cognitive load in recognizing the rolling direction of the interface design with symbol 1 (Table 3).

The results show that the average blink rate and average pupil diameter of participants were higher in the original interface compared to the interface with increased font sizes when identifying the cognitive task corresponding flight state. When identifying the airspeed of the interface, the average blink rate of all valid participants was highest after the font size was increased by 20%, and the average pupil diameter was largest in the original interface. When considering the height of the interface, both the average blink rate and average pupil diameter were highest in the original interface. Meanwhile, when considering the heading of the interface, the average blink rate was highest in the interface with 50% increase in font size. The study found that the average pupil diameter was smallest in the interface with

Table 4. *Mean values of effective feature metrics were calculated for the interface with different font sizes and cognitive tasks*

		Flight state	Airspeed	Altitude	Heading
Accuracy (%)	Original interface	77.78	**97.22**	**86.11**	72.22
	20% increase in font size	77.78	66.67	77.78	**83.33**
	50% increase in font size	**88.89**	88.89	72.22	66.67
reaction time (s)	Original interface	4.0912	3.9217	5.1716	7.5323
	20% increase in font size	**3.3539**	5.9515	4.1216	**4.1880**
	50% increase in font size	3.7756	**3.9143**	**3.2270**	4.5616
Parietal Alpha PSD ($\times 10^{-10}$ μV^2/Hz)	Original interface	3.9987	3.5591	3.5481	3.8637
	20% increase in font size	3.4978	4.3068	4.0015	4.9499
	50% increase in font size	**5.2171**	**5.5716**	**4.9634**	**5.5732**
Frontal Theta PSD ($\times 10^{-8}$ μV^2/Hz)	Original interface	**4.5312**	5.4207	5.8944	4.7266
	20% increase in font size	4.5595	4.9094	**4.2222**	5.2758
	50% increase in font size	4.6536	**4.9092**	5.4941	**4.3560**
Blink rate (bps)	Original interface	**0.5707**	0.6516	**0.6875**	0.6446
	20% increase in font size	0.4699	**0.6852**	0.5434	0.5539
	50% increase in font size	0.5413	0.5565	0.5157	**0.6786**
Pupil diameter (mm)	Original interface	3.8052	3.8624	3.7723	3.8318
	20% increase in font size	3.7869	3.7640	**3.6971**	3.7894
	50% increase in font size	**3.6815**	**3.7103**	3.7654	**3.7117**

*bold text indicates that these evaluation metrics performed relatively better when participants read the corresponding interface.

50% increase in font size and decreased as the font size increased. This suggests that larger font size may lead to a relatively lower cognitive load for the heading cognition task (Table 4).

3.4 EEG frequency features

Figure 7 shows the frequency domain characteristics of the EEG signals of all participants during different cognitive tasks under different interface designs. Some invalid data were excluded due to poor electrode contact.

From the results, it was found that the average power spectral density of the parietal Alpha band was higher in the interface with original symbol and symbol 3 than another two symbols in most of cognitive tasks about the different aircraft symbols, while the average power spectral density of the frontal Theta band was lower in the interface with original symbol and symbol 2 (see Table 3).

In all cognitive tasks of the interface designs with different font sizes, the average power spectral density of the parietal alpha band was higher in the interface with 50% increase in font size than in the other two interfaces. Additionally, the average power spectral density of the parietal alpha band increased with font enlargement, except for the cognition of the flight state in the interface with 20% increase in font size. In contrast, the changes in the average power spectral density of the frontal theta band exhibited minimal variability across cognitive tasks and interface designs, with no discernible pattern (see Table 4).

4.0 Discussion

Due to individual differences and the random sequence of cognitive tasks, each participant may exhibit differences in cognitive performance, cognitive load, and attention levels when faced with different interface designs. Moreover, based on the experimental results, it was found that there were some differences in the performance of each feature indicator for different cognitive tasks in the same interface. Therefore,

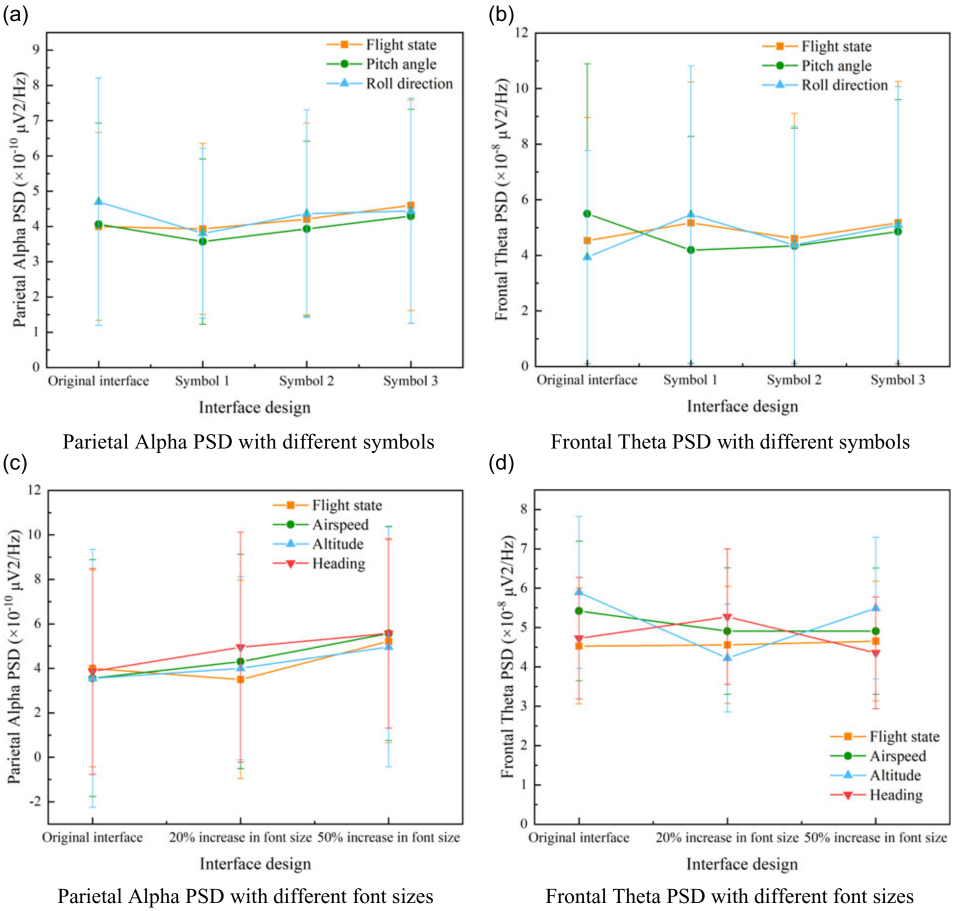


Figure 7. *The differences in parietal alpha PSD and pupil diameter across cognitive tasks among six PFD interface with different aircraft symbols and font sizes.*

it is not possible to determine which PFD interface design is superior, with lower cognitive load, or better cognitive effects based solely on average values.

In this section, we conducted a comprehensive assessment by combining the number of participants occupying the proportion of all valid participants in the change trend of each assessment index under different interface designs. As accuracy increased, reaction time decreased, parietal Alpha PSD increased, frontal Theta PSD decreased, blink rate increased, and pupil diameter decreased, participants' task performance improved and cognitive load reduced. This allowed us to select the interface design that performed relatively better among all interface designs with different aircraft symbols and font sizes (see Tables 5 and 6). In Tables 5 and 6, bold text indicates that the number of participants showing improved performance in these metrics for a particular interface design exceeds 50% of the total participants, signifying that this interface design is superior.

4.1 Discussion of the interface with different aircraft symbols

Combining Tables 2, 3 and 5, it can be observed that among the PFD interfaces with different symbols, the interface with symbol 1 exhibits better cognitive performance in tasks related to recognizing

Table 5. *Percentage changes in evaluation index trends for different aircraft symbols*

	Task	Accuracy increased	Reaction time decreased	Parietal Alpha PSD increased	Frontal Theta PSD decreased	Blink rate increased	Pupil diameter decreased
Symbol 1 compare to original interface	Flight state	**55.56%**	44.44%	**66.67%**	**55.56%**	44.44%	22.22%
	Pitch angle	**77.78%**	**55.56%**	44.44%	33.33%	**66.67%**	22.22%
	Roll direction	**100%**	**66.67%**	**55.56%**	33.33%	**66.67%**	44.44%
Symbol 2 compare to original interface	Flight state	33.33%	44.44%	**77.78%**	**66.67%**	**55.56%**	11.11%
	Pitch angle	44.44%	44.44%	**55.56%**	**55.56%**	44.44%	33.33%
	Roll direction	**88.89%**	44.44%	**55.56%**	44.44%	33.33%	33.33%
Symbol 3 compare to original interface	Flight state	33.33%	33.33%	**77.78%**	**66.67%**	44.44%	**66.67%**
	Pitch angle	**55.56%**	**66.67%**	**66.67%**	44.44%	**55.56%**	**66.67%**
	Roll direction	**55.56%**	**77.78%**	33.33%	22.22%	22.22%	**66.67%**
Symbol 2 compare to Symbol 1	Flight state	44.44%	**55.56%**	33.33%	**55.56%**	44.44%	33.33%
	Pitch angle	11.11%	**55.56%**	**66.67%**	44.44%	44.44%	33.33%
	Roll direction	**50%**	22.22%	44.44%	**55.56%**	33.33%	44.44%
Symbol 3 compare to Symbol 1	Flight state	11.11%	33.33%	55.56%	**66.67%**	44.44%	**66.67%**
	Pitch angle	0	44.44%	**77.78%**	44.44%	33.33%	**66.67%**
	Roll direction	0	33.33%	33.33%	**55.56%**	11.11%	**77.78%**
Symbol 3 compare to Symbol 2	Flight state	44.44%	44.44%	**55.56%**	**66.67%**	33.33%	**77.78%**
	Pitch angle	**66.67%**	**77.78%**	44.44%	33.33%	44.44%	**55.56%**
	Roll direction	33.33%	**55.56%**	44.44%	22.22%	**66.67%**	**77.78%**

Table 6. Percentage changes in evaluation index trends for different font sizes

	Task	Accuracy increased	Reaction time decreased	Parietal Alpha PSD increased	Frontal Theta PSD decreased	Blink rate increased	Pupil diameter decreased
20% increase in font size compare to original interface	Flight state	**66.67%**	**66.67%**	33.33%	**55.56%**	**55.56%**	**55.56%**
	Airspeed	44.44%	33.33%	44.44%	**66.67%**	**55.56%**	**55.56%**
	Altitude	**66.67%**	**66.67%**	**55.56%**	44.44%	44.44%	**55.56%**
	Heading	**88.89%**	**77.78%**	**55.56%**	22.22%	**66.67%**	44.44%
50% increase in font size compare to original interface	Flight state	**88.89%**	**66.67%**	**77.78%**	**66.67%**	**55.56%**	**77.78%**
	Airspeed	**88.89%**	**66.67%**	**55.56%**	**55.56%**	33.33%	**66.67%**
	Altitude	**55.56%**	**55.56%**	**66.67%**	44.44%	44.44%	**55.56%**
	Heading	**66.67%**	**66.67%**	**66.67%**	**66.67%**	**55.56%**	**77.78%**
50% increase in font size compare to 20% increase	Flight state	**66.67%**	44.44%	**66.67%**	**55.56%**	**55.56%**	**77.78%**
	Airspeed	**66.67%**	**66.67%**	**77.78%**	**66.67%**	33.33%	**66.67%**
	Altitude	33.33%	**77.78%**	**55.56%**	33.33%	44.44%	33.33%
	Heading	33.33%	**77.78%**	**66.67%**	**77.78%**	**55.56%**	**66.67%**

roll direction, while the original interface shows relatively lower cognitive load, and its cognitive performance is comparable to that of symbol 1. In tasks related to recognizing pitch angle, although the interface with symbol 2 has lower cognitive load compared to other interface designs, its cognitive performance is poorer and more prone to misleading information. The cognitive effects of the interfaces with other three symbols are similar. In tasks related to recognizing flight state, the cognitive performance and cognitive load of the original interface are generally better than those of the other three symbols, indicating relatively superior cognitive effects.

Therefore, among the various interface designs with different aircraft symbol, the original interface is the most conducive to aircraft flight, as it ensures both cognitive effect and cognitive load are not excessive.

4.2 Discussion of the interface with different font sizes

Combining Tables 2, 4 and 6, it was found that most participants performed better with the interface with 50% font size enlargement compared to the original interface design and the interface with 20% font size enlargement. The interface with 50% font enlargement achieved greater improvements in accuracy, reaction time, and brain load reduction. The trend of improved cognitive performance and reduced cognitive load was particularly evident in the flight state and heading cognition tasks. However, the performance of the interface in airspeed and altitude cognition tasks is inferior to that of the original interface and the interface with a 20% font size enlargement. This issue may be caused by oversized fonts, which can negatively impact the aesthetics and integrity of the data displayed on the airspeed and altitude indicators, leading to relatively poor readability and a poor cognitive effect.

Therefore, through optimizing the font size of the original PFD interface, we can appropriately increase the font size while ensuring the harmony and practicality of the displayed information, to maintain the readability of the PFD interface. This will increase the search efficiency and accuracy of the information displayed on the PFD interface, resulting in a PFD interface that reduces the cognitive load of pilots and ensures flight safety.

4.3 Limitation

In this study, we conducted interface cognition experiments on a computer, investigating the preliminary effects of interface symbols and font sizes on pilot visual cognition through various objective and subjective physiological indicators. We conducted initial screening to optimize the interface design and identify relatively superior designs.

However, compared to dynamically experiencing various cockpit PFD interfaces in a simulated cockpit, this study has limitations. Experimental outcomes may differ from those obtained in dynamic experiments conducted in a simulated cockpit, where participants experience greater immersion. Task performance and workload may vary accordingly, leading to potentially different impacts of different cockpit PFD interface designs on participants compared to the current static experiment. This is a limitation of our current research. In future studies, we plan to integrate selected optimized interfaces into simulated cockpits and conduct dynamic experiments. We will employ visual cognition or attention assessment models to gain deeper insights into how these design factors affect pilot visual cognition and task performance.

5.0 Conclusion

This study is oriented to the optimization design of the cockpit main flight display interface. It optimizes the symbols reflecting the aircraft's flight status and the interface font size, resulting in three optimized symbol interface designs and two optimized font interface designs. In addition, the ergonomics of the PFD interfaces before and after optimization are verified through interface cognition experiments. The results indicate that only the interface with original symbol design can reduce the cognitive load while ensuring task performance, which is more conducive to pilots' perception and comprehension of the aircraft's flight status. As for the interface design with increased font size, it can be observed that the cognitive performance is enhanced with increased font size, while the cognitive load is relatively lower. Furthermore, the ergonomics of the interface with increased font size is superior to that of the original PFD interface design.

This study provides experimental support and optimization suggestions for the optimal design of cockpit PFD interfaces. When optimizing PFD interface design, it is advisable to select symbols that are prominently shaped yet harmonious with the overall interface, colored to distinguish from the background without appearing abrupt. Additionally, the font size should be moderate, balancing readability without overwhelming or impairing pilot cognition.

However, this study also has certain limitations, such as incomplete coverage of design factors and a relatively limited number of participants. Conducting static experiments on a computer may not fully capture the dynamics observed in simulated cockpit experiments. Future studies can further explore the effects of various design factors on pilots' visual cognition, such as color, layout, etc. In addition, these design factors can be combined and optimized to enhance the visual comfort and cognitive effects of the PFD interface, thereby further promoting the improvement and optimization of the cockpit PFD interface. This aims to design cockpit display interfaces that are more compatible with the cognitive characteristics of pilots, thereby enhancing their cognitive efficiency and flight safety.

Acknowledgments.. This paper was financially supported by the Joint Fund of National Natural Science Foundation of China and Civil Aviation Administration of China (No. U2033202, U1333119); the National Natural Science Foundation of China (No.52172387); the Fundamental Research Funds for the Central Universities (ILA22032-1A); and the Aeronautical Science Foundation of China (2022Z071052001); and the Postgraduate Research & Practice Innovation Program of NUAA (xcxjh20230729).

References

[1] Carroll, M. and Dahlstrom, N. Human computer interaction on the modern flight deck, *Int. J. Human–Comput. Interact.*, 2021, **37**, (7), pp 585–587.

[2] Källström, J., Granlund, R. and Heintz, F. Design of simulation-based pilot training systems using machine learning agents, *Aeronaut. J.*, 2022, **126**, (1300), pp 907–931. https://doi.org/10.1017/aer.2022.8

[3] Wang, X., Guo, W., Zhong, Z., Zeng, R., Zhang, J. and Wang, L. The research of touch screen usability in civil aircraft cockpit, *Plos One*, 2024, **19**, (2), p e0292849.

[4] Yu, W., Jin, D., Zhao, F. and Zhang, X. Towards pilot's situation awareness enhancement: A framework of adaptive interaction system and its realization, *ISA Trans.*, 2023, **132**, pp 109–119.

[5] Wei, H., Zhuang, D., Wanyan, X. and Wang, Q. An experimental analysis of situation awareness for cockpit display interface evaluation based on flight simulation, *Chin. J. Aeronaut.*, 2013, **26**, (4), pp 884–889.

[6] Li, W.C., Zakarija, M., Yu, C.S. and McCarty, P. Interface design on cabin pressurization system affecting pilot's situation awareness: The comparison between digital displays and pointed displays, *Hum. Factors Ergon. Manuf. Serv. Ind.*, 2020, **30**, (2), pp 103–113.

[7] Chaparro, A., Miranda, A. and Grubb, J. Aviation displays: Design for automation and new display formats, *Human Fact. Aviat. Aerosp.*, 2023, pp 341–371.

[8] Lanoix, C.A., Rawal, S. and Doyon-Poulin, P. Mechanical device or touchscreen widget: The effects of input device and task size on data entry on the primary flight display, *Int. J. Human–Comput. Interact.*, 2023, pp 1–17.

[9] Tan, W., Wang, W. and Sun, Y. The evaluation model of pilot visual search with onboard context-sensitive information system, In *International Conference on Human-Computer Interaction* (pp. 238–252). Cham: Springer Nature Switzerland, 2023.

[10] Wei, Z., Zhuang, D., Wanyan, X., Liu, C. and Zhuang, H. A model for discrimination and prediction of mental workload of aircraft cockpit display interface, *Chin. J. Aeronaut.*, 2014, **27**, (5), pp 1070–1077.

[11] Taheri Gorji, H., Wilson, N., VanBree, J., Hoffmann, B., Petros, T. and Tavakolian, K. Using machine learning methods and EEG to discriminate aircraft pilot cognitive workload during flight, *Sci. Rep.*, 2023, **13**, (1), p 2507.

[12] Haslbeck, A., Kirchner, P., Schubert, E. and Bengler, K. A flight simulator study to evaluate manual flying skills of airline pilots, In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2014, **58**, (1), pp 11–15). Sage CA: Los Angeles, CA: SAGE Publications.

[13] Zhou, B., Ding, L., Chen, B., Shi, H., Ao, Y., Xu, R. and Li, Y. Physiological characteristics and operational performance of pilots in the high temperature and humidity fighter cockpit environments, *Sensors*, 2021, **21**, (17), p 5798.

[14] Li, W.C., Yu, C.S., Greaves, M. and Braithwaite, G. How cockpit design impacts pilots' attention distribution and perceived workload during aiming a stationary target, *Proc. Manuf.*, 2015, **3**, pp 5663–5669.

[15] Chen, H., Liu, S., Wanyan, X., Pang, L., Dang, Y., Zhu, K. and Yu, X. Influencing factors of novice pilot SA based on DEMATEL-AISM method: From pilots' view, *Heliyon,* 2023, **9**, (2), p. e13425.

[16] Wu, J., Li, C. and Xue, C. How information access, information volume of head-up display and work experience affect pilots' mental workload during flight: An EEG study, In I*nternational Conference on Human-Computer Interaction* (pp. 154–167). Cham: Springer Nature Switzerland, 2023.

[17] Nagasawa, T. and Li, W.C. The application of a human-centred design integrated touchscreen display with automated electronic checklist in the flight deck, In International Conference on Human-Computer Interaction (pp. 135–149). Cham: Springer Nature Switzerland, 2023.

[18] Liang, C., Shuang, L.I.U., Wanyan, X., Chengping, L.I.U., Xu, X.I.A.O. and Yuchen, M.I.N. Effects of input method and display mode of situation map on early warning aircraft reconnaissance task performance with different information complexities, *Chin. J. Aeronaut.*, 2023, **36**, (1), pp 105–114.

[19] Li, W.C., Zhang, J., Le Minh, T., Cao, J. and Wang, L. Visual scan patterns reflect to human-computer interactions on processing different types of messages in the flight deck, *Int. J. Ind. Ergon.*, 2019, **72**, pp 54–60.

[20] Li, W.C., Horn, A., Sun, Z., Zhang, J. and Braithwaite, G. Augmented visualization cues on primary flight display facilitating pilot's monitoring performance, *Int. J. Hum. Comput. Stud.*, 2020, **135**, p 102377.

[21] Şenol, M.B. A new optimization model for design of traditional cockpit interfaces. *Aircr. Eng. Aerosp. Technol.*, 2020, **92**, (3), pp 404–417.

[22] Zhang, X., Cheng, J.A., Xue, H. and Chen, S. Interface design of head-worn display application on condition monitoring in aviation, *Sensors*, 2023, **23**, (2), p 736.

[23] Valverde, R. *Principles of Human Computer Interaction Design: HCI Design*. LAP Lambert Academic Publishing, 2011.

[24] Pandey, A., Panday, S.P. and Joshi, B. Design and development of applications using human-computer interaction, In *Innovations in Artificial Intelligence and Human-Computer Interaction in the Digital Era* (pp 255–293). Academic Press, 2023.

[25] Prati, E., Villani, V., Grandi, F., Peruzzini, M. and Sabattini, L. Use of interaction design methodologies for human–robot collaboration in industrial scenarios, *IEEE Trans. Autom. Sci. Eng.*, 2021, **19**, (4), pp 3126–3138.

[26] Pan, S. Design of intelligent robot control system based on human–computer interaction, *Int. J. Syst. Assur. Eng. Manag.*, 2023, **14**, (2), pp 558–567.

[27] Shen, Z., Zhang, L., Li, R., Hou, J., Liu, C. and Hu, W. The effects of color combinations, luminance contrast, and area ratio on icon visual search performance, *Displays*, 2021, **67**, p 101999.

[28] Rettenmaier, M., Schulze, J. and Bengler, K. How much space is required? Effect of distance, content, and color on external human–machine interface size, *Information*, 2020, **11**, p 7.

[29] Yang, H., Zhang, J., Wang, Y. and Jia, R. Exploring relationships between design features and system usability of intelligent car human–machine interface, *Rob. Auton. Syst.*, 2021, **143**, p 103829.

[30] Dou, J., Xu, C., Chen, S., Xue, C. and Li, X. AR HUD interface optimization model for balancing driver's visual sensitivity and fatigue, *Proc. Comput. Sci.*, 2022, **214**, pp 1568–1580.

[31] Chao, G. Human-computer interaction: process and principles of human-computer interface design, In *2009 International Conference on Computer and Automation Engineering* (pp. 230–233). IEEE, 2009.

[32] Tan, Z., Dai, N., Su, Y., Zhang, R., Li, Y., Wu, D. and Li, S. Human–machine interaction in intelligent and connected vehicles: A review of status quo, issues, and opportunities, IEEE Trans. Intell. Transp. Syst., 2021, 23, (9), pp 13954–13975.

[33] Yin, Y., Zheng, P., Li, C. and Wang, L. A state-of-the-art survey on augmented reality-assisted digital twin for futuristic human-centric industry transformation, *Rob. Comput. Integr. Manuf.*, 2023, **81**, p 102515.

[34] Hao, L. and Chung, W.J. Human-machine interface visual communication design model of electronic equipment using machine vision technology, *Wirel. Commun. Mobile Comput.*, 2022.

[35] Jinjun, X., Miaomiao, F. and Yumeng, Z. Based on visual cognition characteristic of helmet-mounted displays a summary of the interface design. In *2021 26th International Conference on Automation and Computing (ICAC)* (pp. 1–6). IEEE, 2021.

[36] Collinson, R.P.G. Displays and man–machine interaction. In *Introduction to Avionics Systems* (pp. 15–72). Cham: Springer International Publishing, 2023.

[37] Qi, L., Yan, P., Wei, W. and Du, Y. Color matching method of HCI interface design driven by aesthetic perception. *Affect. Pleasur. Des.*, 2022, **41**, p 72.

[38] Faul, F., Erdfelder, E., Lang, A.G. and Buchner, A. G∗ Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences, *Behavior Res. Methods*, 2007, **39**, (2), pp 175–191.

[39] Dattel, A.R., Babin, A.K. and Wang, H. Human factors of flight training and simulation. In *Human Factors in Aviation and Aerospace* (pp. 217–255). Academic Press, 2023.

[40] Puranik, T.G., Rodriguez, N. and Mavris, D.N. Towards online prediction of safety-critical landing metrics in aviation using supervised machine learning, *Transport. Res. C: Emerg. Technol.*, 2020, **120**, p 102819.

[41] Friedrich, M. and Vollrath, M. Human–machine interface design for monitoring safety risks associated with operating small unmanned aircraft systems in urban areas, *Aerospace*, 2021, **8**, (3), p 71.

[42] Melnicuk, V., Thompson, S., Jennings, P. and Birrell, S. Effect of cognitive load on drivers' state and task performance during automated driving: Introducing a novel method for determining stabilisation time following take-over of control, *Accid. Analy. Prevent.*, 2021, **151**, p 105967.

[43] Xiao, Y., Miao, K. and Huang, Y. The effects of graphical encodings on reading performance of the data chart displayed on the periphery of attention, *Displays*, 2023, **77**, p 102378.

[44] Zhang, X., Qu, X., Xue, H., Zhao, H., Li, T. and Tao, D. Modeling pilot mental workload using information theory, *Aeronaut. J.*, 2019, **123**, (1264), pp 828–839. https://doi.org/10.1017/aer.2019.13

[45] Kosch, T., Karolus, J., Zagermann, J., Reiterer, H., Schmidt, A. and Woźniak, P.W. A survey on measuring cognitive workload in human-computer interaction, *ACM Comput. Surv.*, 2023, **55**, (13s), pp 1–39.

[46] Luzzani, G., Buraioli, I., Demarchi, D. and Guglieri, G. 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, *Aeronaut. J.*, 2024, **128**, (1323), pp 928–949. https://doi.org/10.1017/aer.2023.101

[47] Prati, E., Villani, V., Grandi, F., Peruzzini, M. and Sabattini, L. Use of interaction design methodologies for human–robot collaboration in industrial scenarios, *IEEE Trans. Autom. Sci. Eng.*, 2021, **19**, (4), pp 3126–3138.

[48] Brooke, J. SUS-A quick and dirty usability scale, *Usab. Eval. Indus.*, 1996, **189**, (194), pp 4–7.

[49] Lewis, J.R. The system usability scale: past, present, and future, *Int. J. Human–Comput. Interact.*, 2018, **34**, (7), pp 577–590.

[50] Vanneste, P., Raes, A., Morton, J., Bombeke, K., Van Acker, B.B., Larmuseau, C.,, Depaepe, F. and Van den Noortgate, W. Towards measuring cognitive load through multimodal physiological data, *Cognit. Technol. Work*, 2021, **23**, pp 567–585.

[51] He, D., Donmez, B., Liu, C.C. and Plataniotis, K.N. High cognitive load assessment in drivers through wireless electroencephalography and the validation of a modified N-back task, *IEEE Trans. Hum.-Mach. Syst.*, 2019, **49**, (4), pp 362–371.

[52] Ayoub, J., Avetisian, L., Yang, X.J. and Zhou, F. Real-time trust prediction in conditionally automated driving using physiological measures. *IEEE Transactions on Intelligent Transportation Systems*, 2023, **24**(12), 14642–14650.

[53] Liu, Y., Yu, Y., Ye, Z., Li, M., Zhang, Y., Zhou, Z., Hu, D. and Zeng, L.L. Fusion of spatial, temporal, and spectral EEG signatures improves multilevel cognitive load prediction, *IEEE Trans. Hum.-Mach. Syst.*, 2023, **53**, (2), pp 357–366.

[54] Khanam, F., Hossain, A.A. and Ahmad, M. Electroencephalogram-based cognitive load level classification using wavelet decomposition and support vector machine, *Brain-Comput. Interf.*, 2023, **10**, (1), pp 1–15.

[55] McDonnell, A.S., Simmons, T.G., Erickson, G.G., Lohani, M., Cooper, J.M. and Strayer, D.L. This is your brain on autopilot: Neural indices of driver workload and engagement during partial vehicle automation, *Hum. Factors*, 2023, **65**, (7), pp 1435–1450.

[56] Haslbeck, A. and Zhang, B. I spy with my little eye: Analysis of airline pilots' gaze patterns in a manual instrument flight scenario, *Appl. Ergon.*, 2017, **63**, pp 62–71.

[57] Guo, Y., Freer, D., Deligianni, F. and Yang, G.Z. Eye-tracking for performance evaluation and workload estimation in space telerobotic training, *IEEE Trans. Hum.-Mach. Syst.*, 2021, **52**, (1), pp 1–11.

[58] Sáiz-Manzanares, M.C., Marticorena-Sánchez, R., Martin Anton, L.J., González-Díez, I. and Carbonero Martín, M.Á. Using eye tracking technology to analyse cognitive load in multichannel activities in University Students. Int. J. Human–Comput. Interact., 2023, 1–19.

[59] Biondi, F.N., Saberi, B., Graf, F., Cort, J., Pillai, P. and Balasingam, B. Distracted worker: Using pupil size and blink rate to detect cognitive load during manufacturing tasks, *Appl. Ergon.*, 2023, **106**, p 103867.

[60] Blundell, J., Collins, C., Sears, R., Plioutsias, T., Huddlestone, J., Harris, D., Harrison, J., Kershaw, A., Harrison, P. and Lamb, P. Multivariate analysis of gaze behavior and task performance within interface design evaluation, *IEEE Trans. Human-Mach. Syst.*, 2023, **53**, (5), 875–884.

Cite this article: Zeng Z., Sun Y., Liu X., Jie Y. and Zeng Y. (2025). Visual cognition-based optimised design of primary flight displays in cockpits. *The Aeronautical Journal*, **129**, 609–626. https://doi.org/10.1017/aer.2024.103