

Received July 26, 2021, accepted August 18, 2021, date of publication August 27, 2021, date of current version September 10, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3108438

# Developing an Improved ACT-R Model for Pilot Situation Awareness Measurement

HAO CHEN<sup>1</sup>, SHUANG LIU<sup>1</sup>, LIPING PANG<sup>1</sup>, XIAORU WANYAN<sup>1</sup>, AND YUFENG FANG<sup>1,2</sup>

<sup>1</sup>School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China

<sup>2</sup>Aviation Industry Corporation of China, Ltd., Beijing 100028, China

Corresponding authors: Liping Pang (pangliping@buaa.edu.cn) and Xiaoru Wanyan (wanyanxiaoru@buaa.edu.cn)

This work was supported in part by the Joint Program of the National Natural Science Foundation of China and Civil Aviation Administration of China under Grant U1733118, in part by the National Natural Science Foundation of China under Grant 71301005, and in part by the Aeronautical Science Foundation of China under Grant 20181330002.

**ABSTRACT** According to a previously built situation-awareness (SA) model based on attention allocation, with the ACT-R (Adaptive Control of Thought, Rational) theory for analyzing pilot cognitive processes for situation elements, an SA mathematical model was improved to predict pilot SA during exposure to different display interfaces and missions. An indicator-display monitoring task was performed under different experimental conditions for SA model verification, while the SA global assessment technique (SAGAT), performance measures, 10-dimensional SA rating technique (10-D SART), and eye movement measures were adopted to comprehensively assess the operator's SA. The experimental results revealed that theoretical prediction values calculated using the improved SA model were strongly correlated with the operation performance, and thus confirming the model validation. The SAGAT was shown to be a more effective method than SART in this research, and the overall SAGAT accuracy rate, as well as the accuracy response time, are effective indices for SA measurement. Eye-movement indices, such as the fixation/saccade ratio, which corresponds to the mode of information perception and extraction, was examined to be sensitive to operator's SA changes. The Advances of the improved SA model have been achieved in predicting and indicating SA by using human behaviors, including operation performance, SAGAT response behaviors, and visual behaviors. Thereby, it provides a new auxiliary tool for quantitative characterization of pilot's SA during cockpit interface design optimization and ergonomic evaluation.

**INDEX TERMS** ACT-R model, cockpit display interface, cognitive modeling, human-machine interaction, situation awareness.

## I. INTRODUCTION

The pilot-cockpit system is an example of a typical complex human-in-the-loop system. With the development of integrated information and intelligence in modern aircraft, the information flow carried by the flight display and control system has become increasingly diverse and complex. Whether for routine tasks or under special circumstances, the pilot must perform dynamic and repetitive processes of information perception, comprehensive judgment, and decision-making. While performing these tasks, the pilot must constantly maintain situation awareness (SA), which is one of the most important abilities for ensuring flight safety. Aviation accident investigations have indicated that 51.6%

of fatal accidents and 35.1% of non-fatal accidents can be attributed to the failure of decision-making. However, most of these are not a result of decision errors but SA errors [1]. For example, SA may be lost because of a pilot's overreliance on the automation system, which can cause an aircraft with no mechanical faults to be piloted toward the ground, resulting in the fatal occurrence of controlled flight into terrain. SA reflects pilot cognitive status of various factors and conditions that affect aircraft and crew members within a specific situation and time period. Research has indicated that at a higher pilot SA level, more rapid and effective operation can be applied, which is more helpful to ensure flight safety [2], [3].

Among the various SA definitions, Endsley's information process-based three-level model is the most popular. This model describes SA as the perception of task-relevant

The associate editor coordinating the review of this manuscript and approving it for publication was Eunil Park<sup>1</sup>.

elements in the environment (the first level of SA, SA1), the comprehension of the meaning in relation to task goals (the second level of SA, SA2), and the projection of the status in the near future (the third level of SA, SA3) [4], [5]. The SA process can be perceived as cumulative, implying that it is difficult for pilots to cognitively reach SA3 before SA1 and SA2 are achieved. Without good interface design, pilots may stop or be interrupted during information acquisition before they reach SA3 [6]. The defined three-level SA model is important not only because it distinguishes different cognitive processes but also owing to the suggestion that different system designs and training modes are required for dealing with problems caused by multiple SA levels. For example, better alarm mechanism design may be needed when there is a problem with SA1, and better preview display design is more suited to address a problem with SA3. Nowadays, SA is important index to assess design concepts and the technical effectiveness of complex technology systems, which can provide valuable diagnostic data for interactive design [7]–[11].

In recent years, numerous studies have been conducted on SA modeling and experimental measurement in the fields of cognitive psychology and ergonomics [12]–[17]. Early research mainly focused on the formation process analysis of the SA cognitive mechanism and the establishment of SA qualitative models. Accordingly, the development of SA quantitative prediction models has become an important engineering application target for research. Compared with the qualitative model, the quantitative model driven by theory can be used for quantitative analysis and calculation, providing SA prediction and ergonomics evaluation functions [18]–[20]. For example, Entin *et al.* proposed the performance sensitivity model, which suggested that SA insufficiency leads to task performance deterioration, and the total performance deterioration is related to the reduction in the cognitive accuracy of each situation element (SE) [21]. Hooney *et al.* developed a man–machine integration design and analysis human performance model for the improvement of multi-operator SA prediction. This model computes SA as a ratio of the actual SA (the number of SEs that are detected or comprehended) to the optimal SA (the number of SEs that are required or desired to complete the task). High-fidelity flight simulation experiments under different task conditions (aviate, separate, and navigate) were performed for validating the model, and the results indicated that the pilot SA level is sensitive to the display interface design [22]. Wickens *et al.* proposed the attention-situation awareness (A-SA) model, which can be divided into two modules. One governs the allocation of attention to events and channels in the environment (corresponding to SA1), and the second draws an inference or understanding of the current (corresponding to SA2) and future (corresponding to SA3) states of the aircraft within this environment. The A-SA model emphasizes the SA errors related to attention resource allocation. According to a review of SA in flight accidents, there is a close relationship between SA errors and inadequate attention [23]. Additionally, Wickens *et al.* proposed the display formatting

and situation awareness model for predicting which type of aviation display interface features are conducive to improving pilot SA. This model was validated against data from a high-fidelity synthetic vision simulation, and an analysis revealed that the prediction results were strongly correlated with the multitask flight control performance, as well as the traffic awareness response time and accuracy [24]. Adaptive Control of Thought-Rational (ACT-R), which was originally developed by Anderson, is a theory of human cognition and performance as well as a framework for developing computational models of human behavior [25]–[27]. For example, Byrne *et al.* constructed an ACT-R-based computational model of a pilot–aircraft–visual scene–taxiway system, which was used for the diagnosis of possible sources of taxi-driver error. To evaluate the impact of Synthetic Vision System technology on pilot performance in commercial aviation, Byrne *et al.* also developed a pilot–vehicle–airport model combined with ACT-R [28], [29]. By analyzing the theories and research on ACT-R, the driver behavior modeling method in ACT-R was employed by Liu *et al.* to present a highway overtaking behavior model as an example. Driver behavior prediction and model verification both indicated that driver behavior modeling is a flexible and effective method for ensuring traffic safety [30].

Usually, SA models must be verified by combining task scenario simulation and several measurement methods. In a review of multiple SA measurement techniques, subjective measures (e.g., three-dimensional SA rating technique, SA-subjective workload dominance, crew SA, SA supervisory rating form), memory probe measures (e.g., SA global assessment technique), performance measures (e.g., external task performance measures, embedded task performance measures), and physiological measures (e.g., eye-movement tracking, electroencephalograms, event-related potentials, brain mapping, heart rate, electrodermal activity) are all commonly applied [6], [31]–[34]. Using these SA measurements, a large amount of valuable research has been conducted, e.g., comparing the advantages and disadvantages of various SA measurements in different application environments, proposing and testing methods for improving operator SA from different perspectives, using SA as an index for display interface or system design evaluation, and revealing the complex relationships between SA and other factors such as the attention allocation, cognitive load, and automation level. For example, Salmon *et al.* compared two different SA measures (freeze probe recall approach and post-trial subjective rating approach) during a military planning task, and the results indicated that only the participant SA scores derived via the freeze probe recall method produced a statistically significant correlation with the performance [4]. By employing the widely used SA global assessment technique (SAGAT), Endsley *et al.* investigated design methods for improving SA in complex system, e.g., a display model with target orientation, highlighting key features, and parallel processing of multi-display models [35], [36]. Using a high-fidelity flight simulator with a focus on the primary flight display (PFD),

with its route guidance design, Wickens *et al.* studied the effects of tunnel characteristics on the pilot route tracking performance, mental workload, and SA (including traffic and terrain awareness). Traffic awareness was measured by observing the operator accuracy rate and the response time for detecting aircraft symbols, and terrain awareness was measured by asking the operator to point out important terrain directions during the flight simulation freezing period [37]. Chudy *et al.* studied pilot's simultaneous data management ability under stress. The intuitive flight display was designed for improving the pilot's SA and decision-making processes, and its effectiveness was tested via a subjective evaluation based on a flight simulation experiment [38]. For studying SA in the pilot-aircraft system, a human-in-the-loop experiment in the flight simulation environment was performed by Wei *et al.* to assess SA using the SAGAT method and physiological measurement. Accordingly, participant SA was analyzed under different cockpit display interface designs [2]. To develop an event-based method for measuring air traffic controller SA, Yang *et al.* simulated radar control scenarios. SA data based on the event-based method and SART, mental-workload ratings based on the air traffic workload input technique and NASA-TLX, the number of flight conflicts, the actual task load, etc. were applied comprehensively to investigate the sensitivity and validity of the event-based SA assessment technique [39].

The aforementioned studies provide valuable reference for SA modeling and measurement methods. However, most of the previous research focused on proposing a relatively independent concept model from the viewpoint of cognitive psychology, and few of the models have been transformed into their corresponding quantitative calculation models, which limits their applicability in engineering. Additionally, few studies have been performed on SA quantitative prediction models—particularly those covering high-level SA and those directed toward the evaluation and optimization of the cockpit display interface design. In a previous study, the forming process of pilot SA was analyzed in combination with ACT-R cognitive theory, and an SA model covering three levels was proposed on the basis of attention allocation. The theoretical model was validated by several experimental measurements [40], [41]. However, the experimental results indicated that the SA prediction value from the theoretical model was not correlated with the operator behavior performance. As high SA is usually necessary for good operation performance, SA should be partly predicted using the operation performance [42]. In this regard, the previous SA model was improved in the present study, and its correlation with the operation performance of an indicator-monitoring simulation task was analyzed to validate the improved model. Furthermore, the changing trends of several typical evaluation indices were investigated with variations in the SA level.

## II. IMPROVED SA MODEL BASED ON ACT-R THEORY

Human performance modeling is useful for understanding new systems and their effects on human task performance,

without resorting to the more expensive methods of human-in-the-loop experiments or simulations [43]. In the present study, the improved SA model was constructed on the basis of the hierarchical analysis of an SA forming processes combined with the ACT-R theory. FIGURE 1 shows the relationship between the three levels of SA and the ACT-R system [30].

In a certain operation environment, the situation related to the current task can be divided into  $n$  parts labeled  $SE_i$ . Each  $SE_i$  is necessary for supporting an operator task with high performance. It is assumed that the operator's attention resource  $A$  shared by the  $n$  SEs is

$$A = (A_1, A_2, \dots, A_i, \dots, A_n). \quad (1)$$

Generally, information is processed from top-down and bottom-up channels simultaneously. The former is a type of automatic search (or selective control of visual system), while the latter is driven by visual features of  $SE_i$ . The attention resource  $A_i$  allocated to a certain  $SE_i$  can then be expressed by (2) [44], [45].  $\beta_i$  indicates the operator expectation for information, which can be manifested by  $SE_i$ 's occurrence frequency;  $V_i$  indicates  $SE_i$ 's information value or the importance for accomplishing the current task;  $Sa_i$  indicates  $SE_i$ 's salience, which can be manifested by its visual coding; and  $E_i$  indicates the operator effort paid for obtaining the information. Considering the fuzziness and randomness of the human cognitive processing mechanism, the  $V_i$  can be further expressed as  $\partial_i u_i$ , where  $\partial_i$  represents the possibility of the operator potential cognitive availability status for  $SE_i$ , and  $u_i$  is the membership degree normalized into [0,1] based on the cognitive evaluation of  $SE_i$ 's priority [46]. Thus, the attention allocation proportion  $f_i$  of  $SE_i$  can be calculated as "(3)":

$$A_i = \beta_i V_i S a_i E_i^{-1} \quad (2)$$

$$f_i = A_i / \sum_{i=1}^n A_i. \quad (3)$$

In the ACT-R system, the  $SE_i$  of the environment is selectively noticed by the visual module. Assuming that the event of paying attention to  $SE_i$ , which is defined as  $a_i$ , occurs at the current time, the preliminary cognitive processing of  $SE_i$  is performed. The occurrence probability of  $a_i$  can be equivalent to the  $SE_i$ 's attention allocation proportion  $f_i$ , as follows:

$$p(a_i) = f_i. \quad (4)$$

When event  $a_i$  occurs, declarative knowledge is extracted from declarative memory by the buffers to activate the operator's cognition, and the cognitive activation amount  $AC_i$  can be defined as the sum of the base-level activation amount  $AC_{0i}$  and the relevant activation amount  $\sum_j^n W_j S_{ji}$ . The base-level activation amount reflects the general usefulness of  $SE_i$  according to operator's past experience. If the fact of recognizing  $SE_i$  has occurred  $t$  times previously,  $AC_{0i} \approx c + 0.5 \ln t$ , where  $c$  is usually set as 0. The relevant activation amount reflects the  $SE_i$ 's relevance to the current context, where  $W_j$  represents the attention weighting of  $SE_i$ , which

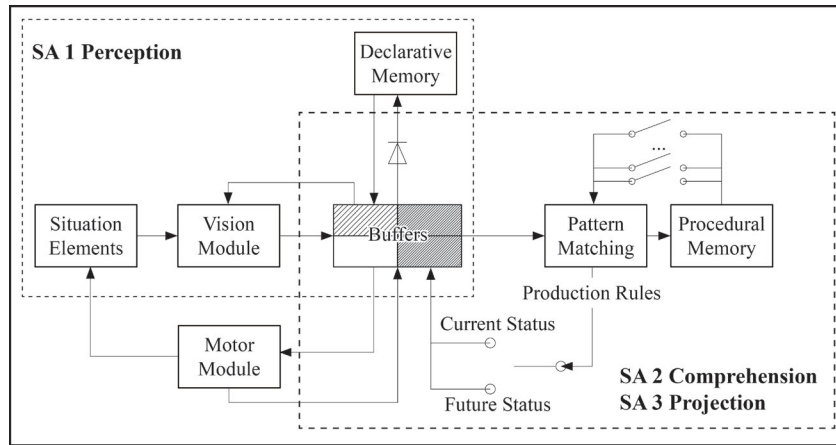


FIGURE 1. Relationship between SA levels and ACT-R system.

can be equivalent to  $f_i$ , and  $S_{ji}$  represents the association intensity between the current  $SE_i$  and the other relational  $SE_j$ . Additionally,  $S_{ji} = S - \ln(fan_j)$ , where  $fan_j$  represents the number of facts associated with  $SE_j$ , and  $S$  is usually set as 2. Only when the cognitive activation amount reaches a certain threshold  $\tau$  can the perception of  $SE_i$  occur (corresponding to SA1). The event of perception of  $SE_i$  is denoted as  $b_i$ . In (5),  $s$  controls the noise at the activation level and is typically set as 0.4, and  $\tau$  is set as 1.0 [25], [46].

$$p(b_i/a_i) = \frac{1}{1 + e^{-(AC_i - \tau)/s}}. \quad (5)$$

According to ACT-R theory, the essence of the cognitive process is the firing of a series of production rules. Multiple production rules can be applied for the “IF-THEN” pattern matching at any point in time, but for recognizing  $SE_i$ , only one production rule with the highest utility  $U_i$  can be selected. After the procedural knowledge is successfully extracted from the procedural memory and optimal pattern matching is performed by the operator cognitive mechanism, the significance of  $SE_i$  can be fully comprehended, which may correspond to SA2 or SA3. In dynamic systems, there is fuzzy boundary between SA2 and SA3 because the understanding of the present usually has direct implications for the future, and both are equally relevant for the task [23], [41]. If the event of  $SE_i$  comprehension is denoted as  $c_i$ , then

$$p(c_i/b_i a_i) = e^{U_i/\theta} / \sum_l^m e^{U_l/\theta}. \quad (6)$$

According to the definition method of the operator cognitive level  $P_i$  for a certain  $SE_i$  [22], there are three situations: the cognitive activation amount is lower than the threshold, the  $SE_i$  is not perceived, and  $P_i$  takes the value of 0; the cognitive activation amount is higher than the threshold, where the declarative knowledge is extracted successfully to form perception, SA1 is achieved, and  $P_i$  takes the value of 0.5; the production rule with the highest utility is matched successfully to form comprehension of the current state or

future state, SA2 or SA3 is achieved, and  $P_i$  takes the value of 1. See (7)–(9):

$$p(a_i \bar{b}_i) = p(a_i)p(\bar{b}_i/a_i) \quad (7)$$

$$p(a_i b_i) = p(a_i)p(b_i/a_i) \quad (8)$$

$$p(a_i b_i c_i) = p(c_i/b_i a_i)p(b_i/a_i)p(a_i). \quad (9)$$

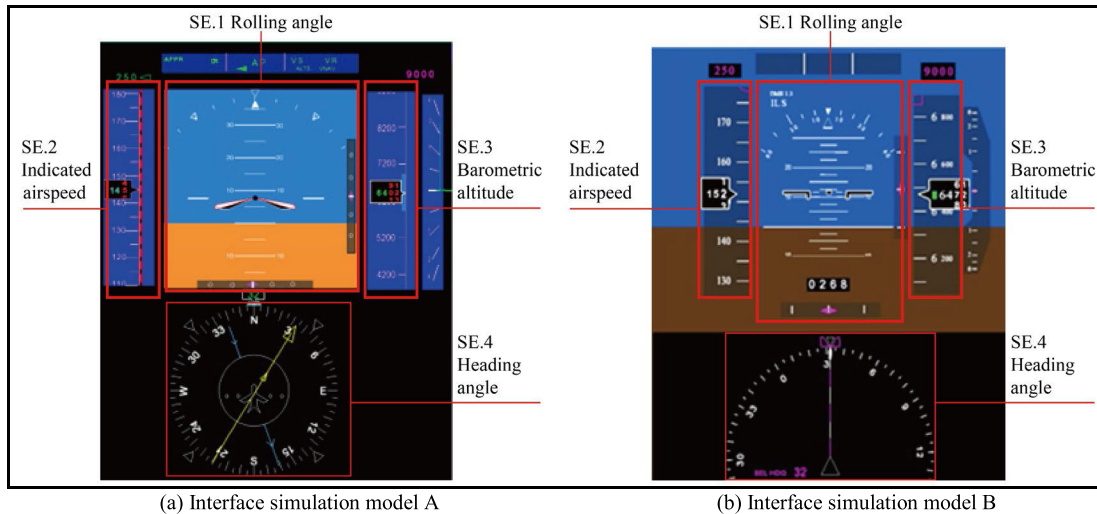
Therefore, the mathematical expectancy  $\bar{P}_i$  of the cognitive level for  $SE_i$  can be calculated as follows [41]:

$$\bar{P}_i = p(a_i b_i c_i) \times 1 + p(a_i b_i) \times 0.5 + p(a_i \bar{b}_i) \times 0. \quad (10)$$

If the operator attempts to obtain and hold a relatively high SA level, the cognitive levels for the critical SEs, which have important impacts on the current task’s performance, should be kept high. Simultaneously, the display attributes of the SEs should contribute to the maintenance of operator SA. Therefore,  $e_i$  is used to represent  $SE_i$ ’s sensitivity coefficient, and its value is related to not only the importance  $u_i$  of  $SE_i$  but also the salient  $Sa_i$  of  $SE_i$ . The relationship between the operator’s whole SA level and the cognitive levels of SEs can then be expressed as follows:

$$\begin{aligned} SA(t_j) &= e_1 \bar{p}_1 + e_2 \bar{p}_2 + \dots + e_n \bar{p}_n \\ &= \sum_{i=1}^n \left( \frac{e^{U_i/\theta}}{\sum_l^m e^{U_l/\theta}} + 0.5 \right) \frac{Sa_i u_i f_i}{1 + e^{-(AC_i - \tau)/s}}. \end{aligned} \quad (11)$$

The SA modeling in our study is focused on the formation process of SA1, SA2, and SA3 on the basis of ACT-R theory. In addition, the SA prediction model is built based on the quantitative representation of pilot’s cognition process rather than simulation calculation. It is conducive to the applicability of the model to flight scenarios, as the model parameters are determined by referring to the literature focusing on pilot cognition and behavior [25], [47]–[50]. However, as the operator SA is affected by multiple factors, the prediction results of our improved SA model cannot represent the absolute-truth value; rather, they represent a value of comparative significance.



**FIGURE 2.** Interface simulation models.

### III. EXPERIMENT METHOD

#### A. PARTICIPANTS

Twenty-eight highly trained, healthy participants (18 males, 10 females) were included in the present study, all from the School of Aeronautic Science and Engineering in Beihang University. All the subjects (aged from 22 to 28 years with a mean age of 23.95 years) were right-handed and possessed normal or corrected-to-normal vision. Written informed consent was obtained from the participants before the experiment.

#### B. INTERFACE SIMULATION MODEL DESIGN

The interface simulation models used in the experiment were designed by referring to two typical interfaces for the PFD, as shown in FIGURE 2. Information layout and color coding differed in the two interface simulation models. GL Studio was adopted as the development platform for graphical modeling of the instruments and control panels, and the C++ programming language and network communication technology were used to realize the virtual instrument simulation programming. The experimental procedure was generated on the Microsoft Visual Studio platform. The interface simulation model was presented on a 17-inch IBM monitor with a resolution of  $1280 \times 1024$ , and the average illumination in the experiment environment was set as 600 lx.

#### C. EXPERIMENT DESIGN

A two-factor, completely within-subjects design was adopted in the experiment. Factor 1 was the probability distribution of indicators' abnormal display, with two levels (level 1: uniform probability distribution, UPD; level 2: non-uniform probability distribution, NUPD) set by the frequency with which SEs were questioned, and factor 2 was the interface simulation model, with two levels (level 1: interface simulation model A; level 2: interface simulation model B).

An indicator-monitoring task was performed in the experiment. The rolling angle (SE.1), indicated airspeed (SE.2), barometric altitude (SE.3), and heading angle (SE.4) presented on the interface simulation model were set as the monitoring targets, referring to the optimal and effective numbers of targets for human attention allocation [40]. Each participant was asked to perform four trials with different combinations of the interface simulation model and probability distributions for indicators' abnormal display. The task orders were counterbalanced across the subjects according to the Latin square design.

#### D. EXPERIMENT PROCEDURE

The SAGAT method was employed in the experiment. For one trial, a total of 32 selection questions (with 4 options for each question) covering three levels of SA were presented randomly on a certain selected interface simulation model according to the chosen probability distribution, so that the participants' cognitive status of all the indicators (SEs) under different task conditions could be obtained. During a random frozen time, only one question aiming at a certain indicator was presented, and the participant responded by selecting one option as soon as possible within the given time limit. The overall accuracy rates for the three levels and the accuracy response time were chosen as the evaluation indices for the SAGAT method. Additionally, combined with the SAGAT method, the participants' performance was recorded simultaneously. Before the experiment, the participants were required to be familiar with the relative importance of each indicator manifested by their scores, which could be obtained by answering the corresponding questions correctly, as well as the probability distributions of indicators' abnormal display, which was manifested by the probability of each indicator being questioned. The participants were required to monitor the interface simulation model on the screen and to capture abnormal information as quickly as possible.

**TABLE 1. Results of improved SA model prediction and experimental measurements (mean  $\pm$  sd).**

	Condition 1 (Model A; UPD)	Condition 2 (Model A; NUPD)	Condition 3 (Model B; UPD)	Condition 4 (Model B; NUPD)
<b>Improved SA model prediction</b>	0.3840	0.4540	0.5205	0.6631
<b>SART Method</b>				
Evaluation scores (points)	19.54 $\pm$ 5.05	20.93 $\pm$ 5.49	19.50 $\pm$ 4.90	20.79 $\pm$ 5.50
<b>SAGAT Method</b>				
Accuracy rate of level 1&2&3	0.71 $\pm$ 0.08	0.73 $\pm$ 0.08	0.72 $\pm$ 0.09	0.74 $\pm$ 0.09
Accuracy response time (s)	2.73 $\pm$ 0.39	2.63 $\pm$ 0.45	2.68 $\pm$ 0.36	2.57 $\pm$ 0.43
<b>Performance</b>				
Operation scores (points)	72.23 $\pm$ 8.43	72.26 $\pm$ 8.83	73.67 $\pm$ 9.37	75.79 $\pm$ 10.27
<b>Eye movement Measurements</b>				
Pupil diameter (mm)	3.60 $\pm$ 0.46	3.53 $\pm$ 0.60	3.59 $\pm$ 0.55	3.96 $\pm$ 1.80
Eyelid opening (mm)	11.20 $\pm$ 1.69	11.16 $\pm$ 1.48	11.25 $\pm$ 1.60	11.15 $\pm$ 1.51
Blink frequency (times/s)	0.32 $\pm$ 0.21	0.34 $\pm$ 0.21	0.32 $\pm$ 0.21	0.36 $\pm$ 0.23
Fixation frequency (times/s)	3.42 $\pm$ 1.37	3.26 $\pm$ 1.03	3.19 $\pm$ 1.15	3.18 $\pm$ 1.01
Fixation duration (s)	333.81 $\pm$ 70.55	316.23 $\pm$ 84.72	329.09 $\pm$ 75.17	319.00 $\pm$ 75.50
Mean fixation time (s)	13.05 $\pm$ 5.78	12.70 $\pm$ 5.70	13.76 $\pm$ 6.25	13.07 $\pm$ 5.75
Mean saccade rate (times/s)	12.77 $\pm$ 1.49	12.84 $\pm$ 1.85	12.61 $\pm$ 1.67	12.61 $\pm$ 1.69
Fixation/Saccade ratio	3.48 $\pm$ 1.69	3.37 $\pm$ 1.61	3.39 $\pm$ 1.70	3.14 $\pm$ 1.48

During each trial, the screen was frozen randomly, and the interface simulation model was immediately replaced by SAGAT question interface on the screen. The question setting of SAGAT was aimed at examining the participants' capability in flight information acquisition, flight status understanding, and near future trends prediction, which corresponds to the three levels of SA (SA1, SA2, and SA3). The participants were asked to click on the correct one of the four options with a mouse within a limited time, clicking on the wrong options or timeout were recorded as an error. Afterwards, the screen freeze was released and the experimental task continued with the previous interface simulation model. All participants were encouraged to obtain scores as high as possible by optimizing their attention allocation strategies. No score was given for an incorrect or missing answer, and the participants' final total scores were used for their performance evaluation. Immediately after each trial, the participants were asked to rate their SA using the 10D-SART (Situation Awareness Rating Technique) self-rating scale [6]. The ratings on 10 individual dimensions were then combined to form a rating for each of the three major factors, including the demand of attention, supply of attention, and understanding of the situation. To objectively record eye-tracking data, the Smart Eye Pro noncontact eye tracker system was employed for real-time tracking during the experiment. The calibration accuracies of all the subjects were always better than  $1^\circ$ , and the sampling rate was 60 Hz.

## IV. RESULTS

### A. MODEL PREDICTION AND EXPERIMENTAL MEASUREMENT RESULTS

The attribute values of 'Effort' and 'Salience' for the four SEs displayed for the two types of interface simulation models were calculated. To be precise, the attribute values of 'Effort' were determined by the relative normalized distances between the SEs, and the attribute values of 'Salience' for each SE were determined by its color matching, indicator

size, and indicator type. Furthermore, the color-matching value was set according to the legibility degree of the displayed characters, the indicator size value was set according to the displayed areas occupied within the experiment interface, and the indicator type value was set according to the difficulty of reading and understanding the indicator [41]. The prediction results for SA calculated via the theoretical model as well as the statistical results for SA obtained from different measuring methods under the four tasks are presented in Table 1.

### B. EXPERIMENTAL ANOVA RESULTS

An analysis of variance (ANOVA) was performed using IBM SPSS 23.0 to investigate the effects of the two factors—the probability distributions of indicators' abnormal display and the interface simulation model—on the participants' SA, operation performance, and eye movement. The level of significance was set at  $\alpha = 0.05$ , yet the significance level lower than 0.1 was recorded as marginally significance for the consideration of the potential research value in the connection between SA prediction model and operator's cognition and behavior. The ANOVA results are presented in Table 2.

The main effect of the interface simulation model on the operation performance was significant ( $F(1, 27) = 5.311$ ,  $p = 0.029$ ,  $\eta^2 = 0.164$ ), as indicated by the higher operation score for interface B compared with interface A. The main effect of the interface simulation model on the SAGAT accuracy rates of level 1 ( $F(1, 27) = 3.309$ ,  $p = 0.080$ ,  $\eta^2 = 0.109$ ) and level 3 ( $F(1, 27) = 3.050$ ,  $p = 0.092$ ,  $\eta^2 = 0.102$ ) was marginally significant. The SAGAT accuracy rates of the two levels in interface B were both higher than those in interface A. In contrast, the main effects of the probability distributions of the indicators' abnormal display, rather than the interface simulation model, on the understanding score in the 10D-SART self-rating scale ( $F(1, 27) = 6.096$ ,  $p = 0.020$ ,  $\eta^2 = 0.184$ ) were significant. The understanding score under the condition of NUPD was higher than that under

TABLE 2. Main effects of experiment anova results.

	Task (UPD & NUPD)			Interface (Model A & Model B)		
	F value	p value	partial $\eta^2$	F value	p value	partial $\eta^2$
<b>SART Method</b>						
SA score_10D-SART	2.375	0.135	0.081	0.013	0.911	0.000
Understanding_10D-SART	6.096	0.020*	0.184	2.460	0.128	0.084
<b>SAGAT Method</b>						
SAGAT accuracy_overall	0.609	0.442	0.022	1.152	0.293	0.041
SAGAT accuracy_level 1	0.517	0.478	0.019	3.309	0.080	0.109
SAGAT accuracy_level 2	1.090	0.306	0.039	0.851	0.364	0.031
SAGAT accuracy_level 3	0.442	0.512	0.016	3.050	0.092	0.102
SAGAT response time	2.318	0.140	0.079	2.543	0.122	0.086
<b>Performance</b>						
Operation score	0.287	0.597	0.011	5.311	0.029*	0.164
Operation accuracy	0.630	0.434	0.023	1.280	0.268	0.045
Operation response time	2.318	0.140	0.079	2.543	0.122	0.086
<b>Eye Movement Measurements</b>						
Fixation frequency	0.517	0.478	0.019	1.398	0.247	0.049
Fixation duration	4.243	0.049*	0.136	0.028	0.869	0.001
Blink frequency	4.933	0.035*	0.154	0.549	0.465	0.020
Blink duration	4.603	0.041*	0.146	2.588	0.119	0.087
Fixation/Saccade ratio	3.499	0.072	0.115	2.469	0.316	0.037
Mean saccade time	0.053	0.820	0.002	3.467	0.074	0.114
Pupil diameter	0.705	0.409	0.025	1.393	0.248	0.049

TABLE 3. Correlations between improved SA model, performance, SAGAT and 10D-SART measurements.

		SA model prediction	Performance (Operation score)	SAGAT accuracy rate (1&2&3)	SAGAT accuracy response time	10D-SART (Evaluation score)
SA model prediction	Correlation	1	0.972(*)	0.943	-0.922	0.452
	Significance		0.028	0.057	0.078	0.548

TABLE 4. Correlations between improved SA model and eye movement measurements.

		SA model prediction	PD	EO	BF	FF	MFT	MSR	F/S ratio
SA model prediction	Correlation	1	0.863	-0.384	0.737	-0.836	0.196	-0.759	-0.955(*)
	Significance		0.137	0.616	0.263	0.164	0.804	0.241	0.045

Note: Pupil diameter = PD, Eyelid opening = EO, Blink frequency = BF, Fixation frequency = FF, Mean fixation time = MFT, Mean saccade rate = MSR, Fixation/Saccade ratio = F/S ratio.

the condition of UPD. No other significant effects on the operation performance or SA were observed.

The main effects of the probability distributions of the indicators' abnormal display on the fixation duration ( $F(1, 27) = 4.243, p = 0.049, \eta^2 = 0.136$ ), blink duration ( $F(1, 27) = 4.603, p = 0.041, \eta^2 = 0.146$ ), and blink frequency ( $F(1, 27) = 4.933, p = 0.035, \eta^2 = 0.154$ ) were significant. The fixation duration under the NUPD condition was lower than that under the UPD condition, while both the blink duration and the blink frequency under the NUPD condition were higher than those under the UPD condition. The main effect of the probability distributions of the indicators' abnormal display on the fixation/saccade ratio ( $F(1, 27) = 3.499, p = 0.072, \eta^2 = 0.115$ ) was also marginally significant, as indicated by the lower fixation/saccade ratio under the NUPD condition compared with the UPD condition. The main effect of the interface simulation model on the mean saccade time ( $F(1, 27) = 3.467, p = 0.074, \eta^2 = 0.114$ ) was marginally

significant, as indicated by the lower mean saccade time in interface B compared with interface A. No other significant effects on the eye movement were observed.

C. CORRELATION ANALYSIS AND MULTILINEAR REGRESSION ANALYSIS RESULTS

To validate the improved SA prediction model based on ACT-R theory, a correlation analysis between the results of the improved SA model prediction and the experimental measurement indices, as well as a multilinear regression analysis, were performed, referring to the method of "modeling the average pilot" [23], [24].

The following conclusions were drawn from the correlation analysis, as shown in Tables 3 and 4. For performance measurements, the improved SA model was highly and significantly correlated with the operation score ( $|r| = 0.972, p = 0.028$ ). For the SAGAT measurements, the improved SA model was highly and close to significantly correlated with

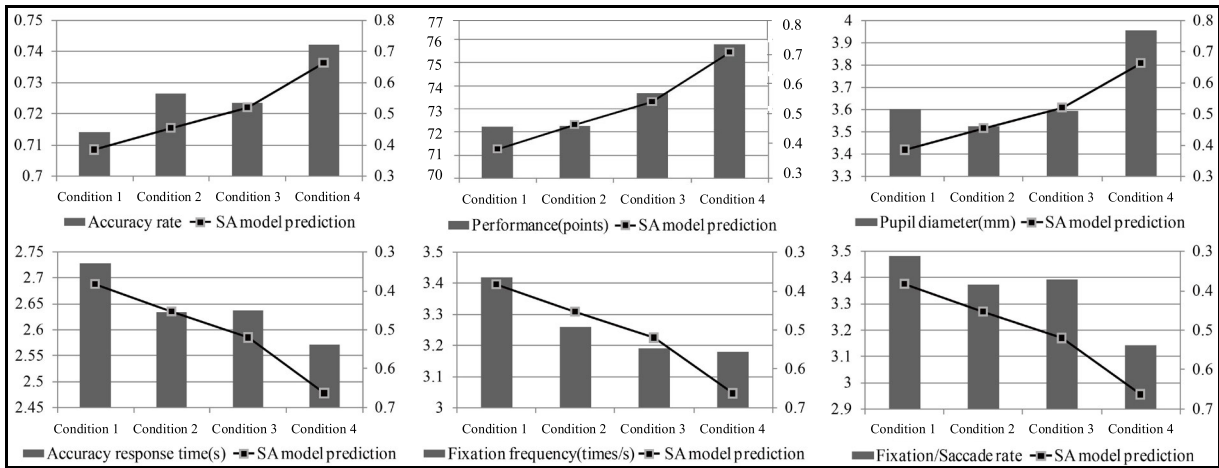


FIGURE 3. Improved SA model prediction results and experimental measurement results under different experimental conditions.

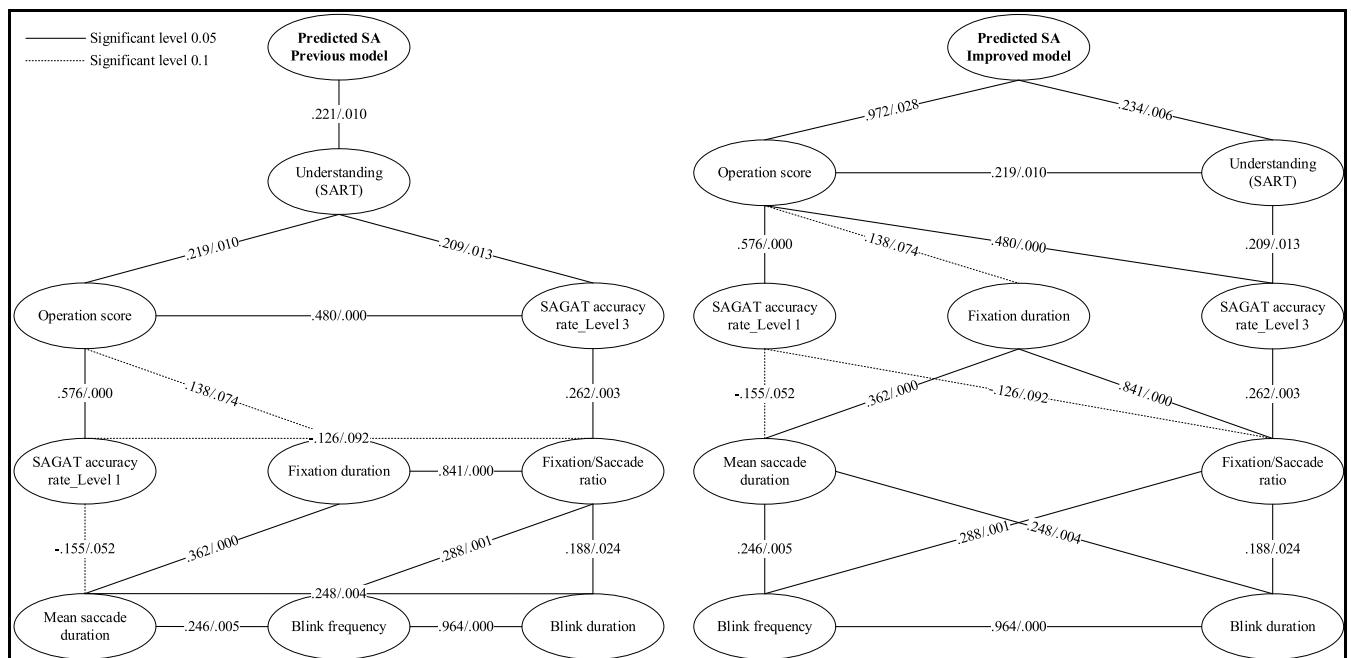


FIGURE 4. Comparison of the previous and improved SA prediction models: correlation between predicted SA and experimental variables with significant effects in the ANOVA.

the overall accuracy rate (level 1&2&3) ( $|r| = 0.943, p = 0.057$ ) and accuracy response time ( $|r| = 0.922, p = 0.078$ ). The improved SA model was not correlated ( $|r| = 0.452, p = 0.548$ ) with the evaluation score for the 10D-SART measurements but was correlated with the dimension of understanding ( $|r| = 0.234, p = 0.006$ ). For the eye-movement measurements, the improved SA model was highly and significantly correlated with the fixation/saccade ratio ( $|r| = 0.955, p = 0.045$ ). It also had high correlations with the pupil diameter ( $|r| = 0.863, p = 0.137$ ) and fixation frequency ( $|r| = 0.836, p = 0.164$ ), but the correlations did not reach significance. As shown in FIGURE 3, the improved SA model prediction

results and experimental measurement results changed under different experimental conditions.

To investigate the consistency between the model prediction and the experimental measurement indices, we analyzed the correlation among the experimental variables with significant effects in the ANOVA, as well as the correlation between the significant variables and the predicted SA in the previous and improved models. The results are shown in FIGURE 4. The operation score was correlated with the predicted SA for the improved model but not for the previous model. The dimension of understanding in 10D-SART was correlated with the predicted SA for both the previous and improved



models. The correlations among the operation score, understanding dimension of 10D-SART, and SAGAT accuracy of level 3 (projection) were significant. The correlation between the SAGAT accuracy of level 1 (perception) and the mean saccade duration was negative, and the correlation between the SAGAT accuracy of level 3 and the fixation/saccade ratio was positive.

Compared with the previous SA model (adjusted  $R^2 = 0.035$ ), the results of the multilinear regression analysis indicated a 14.3% increase in the interpretation rate (adjusted  $R^2$ ) for the improved SA model (adjusted  $R^2 = 0.040$ ). The linear regression equations of the previous and improved SA models are as follows:

$$SA_{\text{previous model}} = 0.331 + 0.008 \cdot \textit{Understanding} \quad (12)$$

$$SA_{\text{improved model}} = 0.326 + 0.008 \cdot \textit{Understanding} \\ + 0.001 \cdot \textit{Performance}. \quad (13)$$

## V. DISCUSSIONS

An improved SA prediction model was proposed with the ACT-R theory, based on the previous one. The model-predicted SA was verified under different experimental conditions by several SA evaluation measurements, including SAGAT, operation performance, 10D-SART, and eye movement measures. The predicted SA value in the improved model has shown more adequate consistency and sensitivity with performance behavior and visual behavior.

The SAGAT is a computerized freeze probe technique. Typically, a simulation task is randomly frozen, all displays are temporarily blanked out, and a series of questions about the current situation at the time of freeze are administered. The operators are required to answer the questions according to their knowledge and judgment. Their situational views are compared with the operational scenario, and an overall SA score is calculated at the end [4]. The SAGAT method is widely used to provide an objective and effective measure of operator SA. In the present study, the SAGAT method was adopted for the assessment of operator SA for different SEs from three levels. As high correlations existed between the theoretical prediction values and accuracy rate, as well as the accuracy response time, according to the experimental results, the improved SA model was validated. Although the accuracy rate of the three levels in the SAGAT method has always been the most widely used and effective index, the accuracy response time is shorter than that used for SA assessment in the previous studies. In the present study, the accuracy response time was highly and negatively correlated with the theoretical prediction value, probably because the SA enhancement can improve the operator cognitive processing speed, reducing the reaction time. Therefore, the accuracy response time can be considered as a good indicator of the SA's changing trend.

There are many studies across different domains in which a high SA has been shown to support good task performance [51], [52]. For example, it is common for researcher to propose an interface alteration that would improve SA and test

it by determining whether the alteration improved the performance [53]. The maintenance of a high SA usually indicates successful operator attention resource allocation strategies [54], which was manifested by the fact that higher scores were obtained by operators in the research. According to the analysis of the experimental results, the theoretical prediction value based on the improved SA model was significantly and positively correlated with the operator performance; thus, the SA model in the present study was confirmed to be optimized compared with the previous model [41].

The 10D-SART self-rating scale was employed in our research to evaluate the operator "demand of attention," "supply of attention," and "understanding of the situation" from 10 different dimensions. However, the experimental results indicated that the correlation between the theoretical prediction value and the evaluation score of the 10D-SART self-rating scale was low, probably owing to the operator's inadequate understanding of the scale and memory decay by post-trial form. Additionally, operators' overestimation or underestimation of their subjective feelings may affect the final evaluation results [55], [56]. Notably, the significant correlation between the operation score and the understanding dimension of 10D-SART, as well as the operation score and the SAGAT accuracy of level 3 (projection) implied the positive effects of comprehension and projection capacity (the two higher levels of SA) on behavioral performance.

In our research, seven eye-movement indices were adopted for SA measurement: the pupil diameter, eyelid opening, blink frequency, fixation frequency, mean fixation time, mean saccade rate, and fixation saccade ratio. Among them, the pupil diameter was highly and positively correlated with SA's theoretical prediction value. This is possibly because when the experimental conditions changed, operators needed to closely monitor the display interface and engage more cognitive efforts (manifesting as the enlargement of pupil diameter) to obtain more information resources [57], so that they could enhance their comprehension and prediction of the dynamically updated information flow, which enhanced their SA accordingly. Additionally, the analysis of the experimental results suggested that a high negative correlation existed between the fixation frequency and the SA's theoretical prediction value. With an increase in the fixation frequency, the consumed time of information extraction per unit time increased, which might indicate the operator's low efficiency in information processing and poor current cognitive level, corresponding to the relatively low SA. Another eye-movement index that was negatively and significantly correlated with the SA's theoretical prediction value in the present study was the fixation/saccade ratio. The fixation/saccade ratio is considered as the ratio of the time spent processing (fixations) information to the time spent searching (saccades) information [58]. Research has indicated that expert pilots exhibit more frequent fixations (viewed as the increment of saccade behaviors) with a shorter dwell time (viewed as the decrement of fixation behaviors) than novice pilots when performing flight simulation tasks [59].

Therefore, a low fixation/saccade ratio may suggest a type of high-efficiency information perception and extraction mode for operators, which results in enhancements in their cognitive states and SA. Unfortunately, the correlations between the other four eye-movement indices (eyelid opening, fixation frequency, mean fixation time, mean saccade rate) and the SA's theoretical prediction value were relatively low. These complex relationships need to be further investigated in future studies.

Compared with the previous SA prediction model [41], the positive correlation coefficient between the predicted SA value and the operation performance was significantly larger for the improved model. This revealed that the human operation performance can be directly indicated by the predicted SA value through the improved model, and vice versa. Similarly, the response behaviors (accuracy rate and response time) in the SAGAT were more closely correlated with the predicted SA in the improved model than in the previous model, confirming the progression of the improved model in measuring the SA according to behaviors. Additionally, the negative correlation between the predicted SA value and the fixation/saccade ratio was far stronger for the improved model than for the previous model, suggesting that the more efficient mode of information perception and extraction (decreased fixation and increased saccade) can be indicated by higher level of the predicted SA value in the improved model.

Summarily, the improved SA prediction model has made progress in the SA being indicated by operator's behaviors, including the operation performance and information processing. The enhanced consistency between the model-predicted SA and the actual performance behavior can contribute to the computability and interpretability of the ACT-R model in analyzing operator's cognition and behavior. Additionally, the consideration of the three levels of SA (i.e., perception, comprehension, and projection) is beneficial for providing more feasible and targeted solutions in aviation engineering (e.g., interface display design, task flow design, and ergonomic evaluation). However, the present study had several limitations. The participants adopted in our experiment were not pilots but graduate students from an aeronautical science and engineering department with experience in operating the simulated flight platform. Considering the participants' ability and experience, we derived simulated flight scenarios from actual flight tasks, the flight task characteristics were properly displayed and the operational procedures were simplified. In addition, the SA prediction model focused more on visual information while insufficiently considered auditory information and communication among stakeholders, which may restrict the model's utility in naturalistic scenarios, such as air route management in air traffic control. Therefore, the reliability and robustness of our conclusion can be further improved by recruiting experienced pilots, and the effects of various typical naturalistic scenarios on the utility of the ACT-R model for predicting human cognition and behavior will be considered in future research.

## VI. CONCLUSION

We improved a previous three-level SA prediction model based on ACT-R cognitive theory and achieved advances in indicating SA by human behaviors, such as the operation performance, SAGAT response behavior, and visual behavior. The support of SA for operation performance can be substantiated in the improved model in view of the variation consistency of SA and performance; thus, the operation performance can be an effective indicator for monitoring and predicting SA through the improved model. The behavioral measurements of SA can be validated by the strengthened correlation between the SAGAT response behaviors and the predicted SA in the improved model. Visual behaviors such as the fixation/saccade ratio can also be an indicator for predicting the SA in the improved model, according to the verified correlation between the SA and the efficiency of the information-processing mode in the improved model. Owing to the substantiated consistency between the model-predicted SA and the performance behavior, as well as the comprehensive consideration of the three levels of SA, the improved SA prediction model provides a new auxiliary tool for quantitative characterization of pilot's SA during cockpit display design optimization and ergonomic evaluation.

## REFERENCES

- [1] H. Jin and L. Wang, "Applying situation awareness to human-machine interface design of aviation," in *Proc. 2nd Int. Symp. Knowl. Acquisition Modeling (KAM)*, vol. 1, 2009, pp. 387–390.
- [2] H. Wei, D. Zhuang, X. Wanyan, and Q. Wang, "An experimental analysis of situation awareness for cockpit display interface evaluation based on flight simulation," *Chin. J. Aeronaut.*, vol. 26, no. 4, pp. 884–889, Aug. 2013.
- [3] Q. Gu, S. Jiang, M. Lian, and C. Lu, "Health and safety situation awareness model and emergency management based on multi-sensor signal fusion," *IEEE Access*, vol. 7, pp. 958–968, 2018.
- [4] P. M. Salmon, N. A. Stanton, G. H. Walker, D. Jenkins, D. Ladva, L. Rafferty, and M. Young, "Measuring situation awareness in complex systems: Comparison of measures study," *Int. J. Ind. Ergonom.*, vol. 39, no. 3, pp. 490–500, May 2009.
- [5] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Hum. Factors*, vol. 37, no. 1, pp. 32–64, Mar. 1995.
- [6] M. R. Endsley, "Theoretical underpinnings of situation awareness: A critical review," in *Situation Awareness Analysis and Measurement*. Mahwah, NJ, USA: Lawrence Erlbaum Associates, 2000, pp. 3–11.
- [7] T. Schnell, Y. Kwon, S. Merchant, and T. Etherington, "Improved flight technical performance in flight decks equipped with synthetic vision information system displays," *Int. J. Aviation Psychol.*, vol. 14, no. 1, pp. 79–102, Feb. 2004.
- [8] M. R. Endsley, "Situation awareness and human error: Designing to support human performance," in *Proc. High Consequence Syst. Surety Conf.*, Albuquerque, NM, USA, 1999, pp. 2–9.
- [9] X. Wang, W. Bo, W. Yang, S. Cui, and P. Chu, "Effect of high-altitude environment on driving safety: A study on drivers' mental workload, situation awareness, and driving behaviour," *J. Adv. Transp.*, vol. 2020, Jul. 2020, Art. no. 7283025.
- [10] Y. Yang, Y. Chen, C. Wu, S. M. Easa, W. Lin, and X. Zheng, "Effect of highway directional signs on driver mental workload and behavior using eye movement and brain wave," *Accident Anal. Prevention*, vol. 146, Oct. 2020, Art. no. 105705.
- [11] Y. Yang, M. Chen, C. Wu, S. M. Easa, and X. Zheng, "Structural equation modeling of drivers' situation awareness considering road and driver factors," *Frontiers Psychol.*, vol. 11, Jul. 2020, Art. no. 1601.
- [12] T. Grundgeiger, A. Hohm, A. Michalek, T. Egenolf, C. Markus, and O. Happel, "The validity of the SEEV model as a process measure of situation awareness: The example of a simulated endotracheal intubation," *Hum. Factors*, early access, Feb. 2021, Art. no. 18720821991651.

- [13] A. O'Brien, G. J. M. Read, and P. M. Salmon, "Situation awareness in multi-agency emergency response: Models, methods and applications," *Int. J. Disaster Risk Reduction*, vol. 48, Sep. 2020, Art. no. 101634.
- [14] C. Wickens, "Attention: Theory, principles, models and applications," *Int. J. Hum.-Comput. Interact.*, vol. 37, no. 5, pp. 403–417, Mar. 2021.
- [15] S. Yan, K. Yao, and C. C. Tran, "Using artificial neural network for predicting and evaluating situation awareness of operator," *IEEE Access*, vol. 9, pp. 20143–20155, 2021.
- [16] Z. Wang, S. Zeng, J. Guo, and H. Che, "A Bayesian network for reliability assessment of man-machine phased-mission system considering the phase dependencies of human cognitive error," *Rel. Eng. Syst. Saf.*, vol. 207, Mar. 2021, Art. no. 107385.
- [17] A. W. Johnson, K. R. Duda, T. B. Sheridan, and C. M. Oman, "A closed-loop model of operator visual attention, situation awareness, and performance across automation mode transitions," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 59, no. 2, pp. 229–241, Mar. 2017.
- [18] J. Rose, C. Bearman, and J. Dorrian, "The low-event task subjective situation awareness (LETSSA) technique: Development and evaluation of a new subjective measure of situation awareness," *Appl. Ergonom.*, vol. 68, pp. 273–282, Apr. 2018.
- [19] Q. Luo, Y. Chen, L. Chen, X. Luo, H. Xia, Y. Zhang, and L. Chen, "Research on situation awareness of airport operation based on Petri nets," *IEEE Access*, vol. 7, pp. 25438–25451, 2019.
- [20] E. Entin and D. Serfaty, "Modeling and measuring situation awareness for target-identification performance," in *Experimental Analysis and Measurement of Situation Awareness*. Daytona Beach, FL, USA: Embry-Riddle Aeronautical University Press, 1995, pp. 233–242.
- [21] B. L. Hooey, B. F. Gore, C. D. Wickens, S. Scott-Nash, C. Socash, E. Salud, and D. C. Foyle, "Modeling pilot situation awareness," in *Human Modelling in Assisted Transportation*. Springer, 2011, pp. 207–213.
- [22] C. D. Wickens, J. McCarley, and L. Thomas, "Attention-situation awareness (A-SA) model of pilot error," in *Proc. NASA Aviation Saf. Program Conf. Hum. Perform. Modeling Approach Landing Augmented Displays*, 2008, pp. 213–239.
- [23] C. D. Wickens, "Display formatting and situation awareness model (DFSAM): An approach to aviation display design," Inst. Aviation, Aviation Hum. Factors Division, Moffett Field, CA, USA, Tech. Rep. AHFD-05-14/NASA-05-5, Oct. 2005.
- [24] M. Fleetwood and M. Byrne, "Modeling the visual search of displays: A revised ACT-R model of icon search based on eye-tracking data," *Hum.-Comput. Interact.*, vol. 21, no. 2, pp. 153–197, May 2006.
- [25] J. Dotlačil, "Building an ACT-R reader for eye-tracking corpus data," *Topics Cogn. Sci.*, vol. 10, no. 1, pp. 144–160, Jan. 2018.
- [26] F. Engelmann, L. A. Jäger, and S. Vasisht, "The effect of prominence and cue association on retrieval processes: A computational account," *Cogn. Sci.*, vol. 43, no. 12, Dec. 2019, Art. no. e12800.
- [27] M. D. Byrne and A. Kirlik, "Using computational cognitive modeling to diagnose possible sources of aviation error," *Int. J. Aviation Psychol.*, vol. 15, no. 2, pp. 135–155, Apr. 2005.
- [28] M. D. Byrne, A. Kirlik, M. D. Fleetwood, D. G. Huss, A. Kosorukoff, R.-S. Lin, and C. S. Fick, "A closed-loop, ACT-R approach to modeling approach and landing with and without synthetic vision system (SVS) technology," in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*. Los Angeles, CA, USA: Sage, 2004, vol. 48, no. 17, pp. 2111–2115.
- [29] Y. Liu and Z. Wu, "Multitasking driver cognitive behavior modeling," in *Proc. 3rd Int. IEEE Conf. Intell. Syst.*, vols. 1–2, Sep. 2006, pp. 49–54.
- [30] P. M. Salmon, N. A. Stanton, and K. L. Young, "Situation awareness on the road: Review, theoretical and methodological issues, and future directions," *Theor. Issues Ergonom. Sci.*, vol. 13, no. 4, pp. 472–492, Jul. 2012.
- [31] T. Lu, Z. Lou, F. Shao, X. You, and M. Tang, "Attention allocation in pilots based on climbing and circling mission behavior," *Psychol. Res.-Psychol. Forschung*, vol. 85, no. 3, pp. 1136–1145, 2021.
- [32] X. Yang and J. H. Kim, "Assessing situation awareness in multitasking supervisory control using success rate of self-terminating search," *Int. J. Ind. Ergonom.*, vol. 72, pp. 354–362, Jul. 2019.
- [33] S. Puuska, L. Rummukainen, J. Timonen, L. Lääperi, M. Klemetti, L. Oksama, and J. Vankka, "Nationwide critical infrastructure monitoring using a common operating picture framework," *Int. J. Crit. Infrastruct. Protection*, vol. 20, pp. 28–47, Mar. 2018.
- [34] M. R. Endsley, "Errors in situation assessment: Implications for system design," in *Human Error and System Design and Management*. London, U.K.: Springer, 2000, pp. 15–26.
- [35] M. Naderpour, J. Lu, and G. Zhang, "A safety-critical decision support system evaluation using situation awareness and workload measures," *Rel. Eng. Syst. Saf.*, vol. 150, pp. 147–159, Jun. 2016.
- [36] M. Martinussen and D. R. Hunter, *Aviation Psychology and Human Factors*. Boca Raton, FL, USA: CRC Press, 2017.
- [37] P. Chudy, K. Rydlo, and T. Konecny, "Intuitive flight display for light aircraft," in *Proc. AIAA Modeling Simulation Technol. Conf.*, Aug. 2011, p. 6348.
- [38] J. Yang, Y. Zeng, K. Zhang, and E. M. Rantanen, "Event-based assessment technique for measuring air traffic controller's situation awareness," *Space Med. Med. Eng.*, vol. 21, no. 4, pp. 321–327, Aug. 2008.
- [39] X. Wanyan, D. Zhuang, H. Wei, and J. Song, "Pilot attention allocation model based on fuzzy theory," *Comput. Math. Appl.*, vol. 62, no. 7, pp. 2727–2735, Oct. 2011.
- [40] S. Liu, X. Wanyan, and D. Zhuang, "Modeling the situation awareness by the analysis of cognitive process," *Bio-Med. Mater. Eng.*, vol. 24, no. 6, pp. 2311–2318, 2014.
- [41] C. D. Wickens, "Situation awareness: Review of Mica Endsley's 1995 articles on situation awareness theory and measurement," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 50, no. 3, pp. 397–403, Jun. 2008.
- [42] J. Keller, K. Leiden, R. Small, A. Goodman, and B. Hooey, "Cognitive task analysis of commercial jet aircraft pilots during instrument approaches for baseline and synthetic vision displays," in *Proc. NASA Aviation Saf. Program Conf. Hum. Perform. Modeling Approach Landing Augmented Displays*, 2003, p. 15.
- [43] C. Wickens, J. McCarley, and K. Steelman-Allen, "NT-SEEV: A model of attention capture and noticing on the flight deck," in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, Oct. 2009. Los Angeles, CA, USA: Sage, vol. 53, no. 12, pp. 769–773.
- [44] C. D. Wickens, J. Goh, J. Helleberg, W. J. Horrey, and D. A. Talleur, "Attentional models of multitask pilot performance using advanced display technology," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 45, no. 3, pp. 360–380, Sep. 2003.
- [45] X. Wu, X. Wanyan, and D. Zhuang, "Pilot's visual attention location modeling under fatigue," *Technol. Health Care*, vol. 23, pp. S373–S381, Jun. 2015.
- [46] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. L. Qin, "An integrated theory of the mind," *Psychol. Rev.*, vol. 111, no. 4, pp. 1036–1060, Oct. 2004.
- [47] D. Gall and T. Frühwirth, "An operational semantics for the cognitive architecture ACT-R and its translation to constraint handling rules," *ACM Trans. Comput. Log.*, vol. 19, no. 3, pp. 1–42, Sep. 2018.
- [48] C. R. Fisher, J. W. Houpt, and G. Gunzelmann, "Developing memory-based models of ACT-R within a statistical framework," *J. Math. Psychol.*, vol. 98, Sep. 2020, Art. no. 102416.
- [49] M. Ilbeygi, M. R. Kangavari, and S. A. Golmohammadi, "Equipping the ACT-R cognitive architecture with a temporal ratio model of memory and using it in a new intelligent adaptive interface," *User Model. User-Adapted Interact.*, vol. 29, no. 5, pp. 943–976, Nov. 2019.
- [50] F. E. Ritter, F. Tehranchi, and J. D. Oury, "ACT-R: A cognitive architecture for modeling cognition," *Wiley Interdiscipl. Rev., Cogn. Sci.*, vol. 10, no. 3, May/Jun. 2019, Art. no. e1488.
- [51] C. M. Perry, M. A. Sheik-Nainar, N. Segall, R. Ma, and D. B. Kaber, "Effects of physical workload on cognitive task performance and situation awareness," *Theor. Issues Ergonom. Sci.*, vol. 9, no. 2, pp. 95–113, Mar. 2008.
- [52] R. Ma and D. B. Kaber, "Situation awareness and workload in driving while using adaptive cruise control and a cell phone," *Int. J. Ind. Ergonom.*, vol. 35, no. 10, pp. 939–953, Oct. 2005.
- [53] M. A. Vidulich, "Testing the sensitivity of situation awareness metrics in interface evaluations," in *Situation Awareness Analysis and Measurement*. Mahwah, NJ, USA: Lawrence Erlbaum Associates, 2000, pp. 227–246.
- [54] R. M. Ratwani, J. M. McCurry, and J. G. Trafton, "Single operator, multiple robots: An eye movement based theoretic model of operator situation awareness," in *Proc. 5th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2010, pp. 235–242.
- [55] M. R. Endsley, S. J. Selcon, T. D. Hardiman, and D. G. Croft, "A comparative analysis of SAGAT and SART for evaluations of situation awareness," in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*. Los Angeles, CA, USA: Sage, 1998, vol. 42, no. 1, pp. 82–86.
- [56] D. G. Jones and M. R. Endsley, "Use of real-time probes for measuring situation awareness," *Int. J. Aviation Psychol.*, vol. 14, no. 4, pp. 343–367, Oct. 2004.

- [57] L. L. Di Stasi, A. Antolí, M. Gea, and J. J. Cañas, “A neuroergonomic approach to evaluating mental workload in hypermedia interactions,” *Int. J. Ind. Ergonom.*, vol. 41, no. 3, pp. 298–304, May 2011.
- [58] J. H. Goldberg and X. P. Kotval, “Computer interface evaluation using eye movements: Methods and constructs,” *Int. J. Ind. Ergonom.*, vol. 24, no. 6, pp. 631–645, Oct. 1999.
- [59] W. Liu, X. Yuan, Z. Liu, W. Kang, and R. Ma, “Experimental study of pilots’ scan and performance, workloads,” *Space Med. Med. Eng.*, vol. 18, no. 4, pp. 293–296, Aug. 2005.



**HAO CHEN** was born in Anhui, China, in 1992. He is currently pursuing the Ph.D. degree in man machine and environment engineering with Beihang University.

He is the author of ten articles, has two inventions, and holds two patents. His research interests include pilot cognition, cockpit interface design, cabin design and optimization, ergonomic evaluation, team collaboration, and work scheduling. He has experience in human cognition and behavioral performance in transportation fields of aviation, navigation, and land.



**SHUANG LIU** was born in Xuancheng, Anhui, China, in 1986. She received the B.S. degree in automation from Anhui University of Science and Technology, Huainan, Anhui, in 2005, and Ph.D. degree in man machine and environment engineering from Beihang University, Beijing, China, in 2015.

From 2015 to 2019, she was a Senior Engineer with the Marine Human Factors Engineering Laboratory, China Institute of Marine Technology and Economy, Beijing. Since 2019, she has been a Postdoctoral Researcher at Beihang University. She is the author of more than 30 articles and has more than ten inventions. Her research interests include the interface design based on human cognitive characteristics, training theory for individual ability and team cooperation ability, and ergonomics evaluation and data analysis technology in complex task environments. She has been in charge of and participated in many important national scientific research projects. She has rich experience in human–computer interface design and engineering application in large ships, aircraft cockpit, armored vehicles, and other fields.



**LIPING PANG** was born in 1973. She received the Ph.D. degree from the School of Aeronautic Science and Engineering, Beihang University, in 2005. Since then, she was an Assistant Professor with the School of Aeronautic Science and Engineering, Beihang University. From 2011 to 2012, she was a Visiting Scholar with the International Centre for Indoor Environment and Energy, Technical University of Denmark. She was promoted to a Professor at Beihang University, in 2014. She is the author of more than 30 articles and five inventions. Her research interests include aircraft environmental control and aviation ergonomics.



**XIAORU WANYAN** was born in Hebei, China, in 1981. She received the B.S. degree in industrial design and the Ph.D. degree in human–machine and environmental engineering from Beihang University, Beijing, China, in 2004 and 2011, respectively.

From 2012 to 2019, she was an Assistant Professor with the School of Aeronautic Science and Engineering, Beihang University, where she has been an Associate Professor, since 2020. She is the author of three books and more than 70 articles. She holds five patents. Her research interests include modeling and measurement of pilot cognitive behavior, ergonomic evaluation, and design of cockpit human–machine interfaces.



**YUFENG FANG** was born in 1961. From 1982 to 1995, he was a Senior Researcher at Shenyang Aircraft Company, Ltd., where he was a Vice Chief Engineer, from 1995 to 2006. From 2008 to 2010, he was the Vice President at Xi’an Aircraft Industry Group Company, Ltd. From 2012 to 2019, he was the Chairman of the board in the entity. Since 2019, he has been a member of the management at Aviation Industry Corporation of China, Ltd. His research interests include aircraft system design, engine and environmental control system design, and human-system integration design.

...