

Mental workload prediction based on attentional resource allocation and information processing

Xu Xiao, Xiaoru Wanyan* and Damin Zhuang

School of Aeronautics Science and Engineering, Beihang University, No. 37 Xueyuan Road, Haidian District, Beijing 100191, China

Abstract. Mental workload is an important component in complex human-machine systems. The limited applicability of empirical workload measures produces the need for workload modeling and prediction methods. In the present study, a mental workload prediction model is built on the basis of attentional resource allocation and information processing to ensure pilots' accuracy and speed in understanding large amounts of flight information on the cockpit display interface. Validation with an empirical study of an abnormal attitude recovery task showed that this model's prediction of mental workload highly correlated with experimental results. This mental workload prediction model provides a new tool for optimizing human factors interface design and reducing human errors.

Keywords: Mental workload, prediction model, attention allocation, information theory, ergonomics

1. Introduction

With the rapid development of intelligent technology and flight automation, a pilot's number of manual control tasks reduces while the number of monitoring control tasks increases. Whether the pilot can perform the flight task successfully largely depends on monitoring a plurality of flight indicators and understanding flight information correctly. Fitts, et al. found that 18% of flight accidents derived from pilots' errors in comprehensive interpretation of flight indicator information [1]. Therefore, researchers propose the mental workload (MW) concept to optimize the display number and coding method of flight information on the cockpit display interface, which will decrease human error and increase system reliability and operator comfort [2]. Ever since MW was first defined in the 1970s [1], researchers have expatiated the effects of MW on humans in dynamic systems through different angles. Parasuraman, et al. defined MW as the functional relationship between the mental resources demanded by a task and the available mental resources provided by the operator [2]. The pilot's MW can be measured by various definitions in different angles. By now, plenty of studies belong to empirical workload measures. The disadvantage is that they discover the design problem

* Address for correspondence: Xiaoru Wanyan, School of Aeronautics Science and Engineering, Beihang University, No. 37 Xueyuan Road, Haidian District, Beijing 100191, China. Tel.: +8601082338163; Fax: +8601082338163; E-mail: wanyanxiaoru@buaa.edu.cn.

after the fact, and there is a high economic cost [1]. Prior analysis of the operator's MW in the human-machine system can be adopted in both MW prediction and evaluation. The widely used MW prediction modeling methods include task network analysis and cognitive structure analysis [3].

Task network analysis does not generally need special mathematical modeling technology, but instead requires detailed task sequence and demanded time information [3]. For example, the Timeline Analysis and Prediction (TLAP) method worked out the timeline through elaborate task analysis and calculated MW in a certain time unit, which was the TR/TA (Time Required/Time Available) in the time unit [4]. Wichens explained the workload problem by the resources' volume and proposed a 4-dimensional multiple resources model. An increase in task difficulty or resource shortage affected by resource conflicts between many tasks caused a decrease in the total performance [5]. Colombi, et al. synthesized a system frame, operator task analysis and workload theory based on the 4-dimensional multiple resources model, so that the operator's MW could be assessed in the early design of Systems of Systems (SoS) [6]. Liang, et al. combined task analysis with McCrachen-Aldrich technology and divided the task into secondary tasks. Then the MW prediction was performed by quantifying the operating procedure from visual, auditory and two other dimensions [7].

Cognitive structure analysis utilizes foundational cognitive theory to establish a frame structure representing a potential object, principle and the operator's performing mechanism; then the human behavior in the complex system can be expressed as relevant factors in the structure [3]. For instance, Xie and Salvendy presented and verified a MW conceptual frame under single or multiple task environments, including instantaneous MW, average MW, accumulated MW and total MW, as well as a prediction model [8]. Cao and Liu conducted a model on the basis of ACTR-QN (integration of Adaptive Control of Thought-Rational and Queueing Network), which presented the multi-dimensional factors of MW, the predict error rate, and total MW [9]. Jo, et al. applied the ACT-R cognitive structure to develop a MW prediction model, which mainly considered the activated time of 8 modules in the ACT-R [10]. The information processing (IP)/perceptual control theory (PCT) model stated that MW consisted of time workload (defined as time pressure) and intensity workload (defined as task difficulty, such as information amount) [4]. Unlike most task network analysis, cognitive structure analysis has no need for detailed task scenarios and task decomposition steps, so it is easier to exploit.

Although the above studies predicted the operator's MW from considerable angles, they did not consider the integrated effects of exogenous task demands and endogenous resource supply on MW. Hence, the present study focused on the MW induced by visual information processing of the aircraft cockpit display interface and developed a MW prediction model by employing cognitive structure analysis and integrating attentional resource allocation and information processing. To verify the validity of this model, an empirical study of an abnormal attitude recovery task under different MW conditions on the display interface was conducted. NASA-TLX subjective evaluation, performance, and eye movement technology were applied to analyze the trend of the subjects' MW and to compare it with the theoretical predictions.

2. Model development

The aircraft cockpit is a special operating environment containing highly dense flight information. 70% of information is delivered visually, so the pilot's MW mainly comes from monitoring visual information tasks. The present study concentrates on modeling and optimizing the MW caused by monitoring information on the cockpit display interface.

With the cockpit display interface divided into n areas of interest (AOI), the total MW equals the addition of MW from n AOIs. In spite of individual differences, MW has two main determinants: exogenous task demands as specified by factors such as task difficulty, accuracy, and time requirement, and endogenous supply of attentional or processing resources such as perceiving, updating memory, and decision making [1]. In the present study, information amount, time pressure, and the integrated performance value of visual coding are applied to quantize exogenous task demands, while the attentional allocation factor quantizes the endogenous resource supply.

The information amount may be adopted to measure task difficulty. The information amount that a pilot completes per unit of time can be applied to evaluate information processing effectiveness, which can be used to compare the different MW that pilots face [4]. Flight information can be quantified by information theory, which states that the information amount is affected by the number of possible events, event probability, and event sequence; as a result, the average information amount appearing with time H_i is shown as Eq. (1) [11]. Shown in Eq. (2), the information probability of AOI P_i represents the information number in AOI i and the coefficient related to the information value. P_{1i} is the value of information in AOI i with respect to all the AOIs. P_{2i} is the SNR of information in AOI i .

$$H_i = P_i [\log_2(1 / P_i)] \quad (1)$$

$$P_i = P_{1i} * P_{2i} \quad (2)$$

Time pressure has been considered to have an important effect on pilot safety concerning information processing. It refers to the mental stress generated when operators feel the impossibility of finishing a task smoothly because of the limited available time [12]. Time pressure can be quantified as the available time set by the system, which may be represented by the operating time from the information presentation to perception, understanding, and finally performance. According to the definition of the rate of information transmission [4], [11], the average information amount can be multiplied by time pressure, which may avoid the quantization problem of the weight coefficient in the linear additive model.

In addition, a series of studies indicate that proper visual coding helps pilot perceive and understand flight information easier, which in turn reduces MW during cognitive processing [13]. The present study establishes a Multi-dimensional Visual Coding Evaluation Model for the cockpit display interface by considering character size, color matching, interface layout, and other visual coding characters that affect MW, taking advantage of meta-analysis and integrated performance values, and involving further statistical analysis of previous experimental results [14]. The integrated performance value C_i of visual coding based on this model is shown in Eq. (3), where w_i^j is the weight coefficient of AOI i , V_i^j is the integrated performance value of visual coding j , and \odot represents the fuzzy weighted averaging operator.

$$C_i = 0.001 \sum_{j=1}^n w_i^j \odot v_i^j \quad (3)$$

As for the endogenous resource supply, attentional resources required by the tasks can be allocated by continuously monitoring visual information spontaneously and independently [15]. In the perception stage, the more attentional resources a pilot allocates on the AOI, the more cognitive resources are occupied, which leads to higher MW. Attentional allocation factor f_i can represent the operator's endogenous MW consumed by paying attention to the AOI. Eq. (4) is the attentional

resource A of n AOIs. Eq. (5) shows that the attentional allocation factor f_i equals the ratio of the attentional resource of AOI $_i$ to the total attentional resource [15]. Besides, to be consistent with MW modeling, the sum of f_i is supposed to be n, where n refers to the number of AOIs.

$$A = (A_1, A_2, \dots, A_i, \dots, A_n) \quad (4)$$

$$f_i = n A_i / \sum_{i=1}^n A_i \quad (5)$$

The visual system captures information and the neural systems transfers it to the brain, so the attentional resource allocated to AOI $_i$ is Eq. (6) [15]. All the information has a probability β_i to be noticed for one fixation. The information will occupy some attentional resources if it is noticed; abnormal information in AOI $_i$ may apply highlighting, flashing, and other display formats to help pilots perceive and respond quickly. Saliency S_i is quantified by display formats; effort E_i refers to the effort operators take to gain information by eye or head movements. It restrains attentional resource allocation to some extent and can be measured by the relative distance between AOIs; V_i is the information importance attribute associated with the task [15], as in Eq. (7). α_i is the probability of the information potential cognitive condition in AOI $_i$, and u_i is the information importance attribute of AOI $_i$. Since there is a certain ambiguity in humans' subjective cognitions, fuzzy entropy h_i is used to reflect the operator's mental activity in the information cognition in AOI $_i$ [15].

$$A_i = \beta_i V_i S_i E_i^{-1} \quad (6)$$

$$V_i = \partial_i u_i = (\exp h_i / \sum_{i=1}^n \exp h_i) u_i \quad (7)$$

$$= \{\exp[-u_i \ln u_i - (1 - u_i) \ln(1 - u_i)] / \sum_{i=1}^n \exp[-u_i \ln u_i - (1 - u_i) \ln(1 - u_i)]\} u_i$$

Starting with exogenous task demand and endogenous resource supply, the above four factors have a corporate affect on the potential MW of a pilot required from the display interface. The MW for the aircraft cockpit display interface is found by putting the calculations of each factor in Eq. (8).

$$MW = \sum_{i=1}^n (H_i * T_i^{-1} * C_i^{-1}) * f_i = \sum_{i=1}^n \{P_i [\log_2(1/P_i)] T_i^{-1} C_i^{-1} (n A_i / \sum_{i=1}^n A_i)\} \quad (8)$$

The integrated performance value of visual coding is $C_i \geq 0$ with unit 1, the average information amount is $H_i \geq 0$ with unit bits, the time pressure is $T_i > 0$ with unit s, and the attentional allocation factors is $f_i \geq 0$ with unit 1. As a consequence, MW is $MW \geq 0$ with unit bits/s.

3. Experimental method

3.1. Subjects

Thirty-two Beihang University graduates (twenty-eight males and four females with a mean age of 24.6) with a background knowledge in aviation participated in the experiment. Each participant was right-handed and possessed normal or corrected to normal vision with no color blindness or weakness. All subjects gave their full informed consent before the experiment.

3.2. Materials

The ergonomics experiment was carried out on the two primary flight display (PFD) interfaces, as shown in Figure 1. According to the practical flight tasks and cockpit PFD interface, the experimental program was established by GL studio virtual instrument simulation software. The experimental interface was presented on a 17-inch LCD monitor with a resolution of 1440×900. During the experiment, environment average illumination remained at about 600 Lx.

The subjects' blink frequency and fixation time were recorded with the non-contact eye tracker system Smart Eye Pro with an accuracy of better than 1° and a 60Hz sampling rate. A 4-point calibration and validation was performed before the start of the each experimental session. Blinks and fixations with durations less than 10 ms and 100 ms, respectively, were not considered in the analysis. Subjects were asked to complete the NASA-TLX [16] subjective questionnaire at the end of the experiment.

3.3. Experiment task

The experiment simulated the abnormal attitude recovery task during take-off, cruise, and landing, which required subjects to continuously monitor numerous information related to flight attitude controlling on the display interface, including pitching angle, rolling angle, air speed, pressure altitude, and heading angle. As soon as the information was beyond the limits set by normal flight conditions, the subject needed to respond within the prescribed time.

3.4. Experiment design

The experiment adopted a 2*2*2 three factors mixed design. The two interfaces (one is interface a,

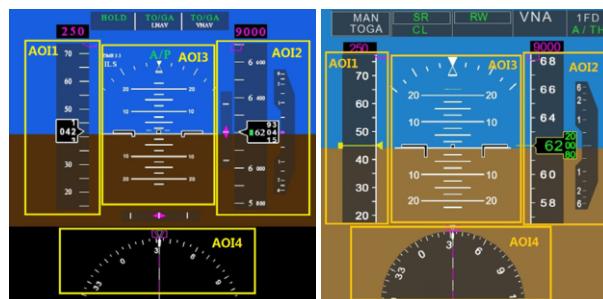


Fig. 1. PFD interface a (left) and PFD interface b (right). Interface a simulated B787 PFD with proper simplification and abstraction, while interface b was the optimized interface based on the multi-dimensional visual coding evaluation model. They were both divided into four AOIs.

the other is interface b), representing different visual coding characters, were between-subjects. The amount of information (two conditions: one monitoring five pieces of information and the second monitoring two pieces of information), standing for the information amount, and the two time conditions (one with a presentation time of 1.5 s and the other with a presentation time of 3 s), standing for the time pressure, were within-subjects. The subjects were randomly assigned into two groups, so that each subject participated in a total of four experiments. The Latin square design was applied to reduce both the practice effect and the fatigue effect.

4. Results

The attentional allocation factors f_i were obtained by calculating the attribute values of effort and salience for four AOIs on two interfaces and integrating probability and the importance of different information. Theoretical prediction MW was figured out in Table 1. The statistic results in Table 1 indicate that a decrease in the amount of information and an increase in time with the optimized interface led to better performance, a lower NASA-TLX subjective score, higher blink frequency, and shorter fixation time. This, in turn, led to a decrease in MW, just as the MW prediction model predicted. Before the experiment, the basal levels of eye movement between the two groups were recorded. And the main effect of between-subject design was insignificant for both blink frequencies ($p=0.952$) and fixation time ($p=0.803$), respectively. The eye movement tracking results were analyzed to get the fixation heat map under two interfaces, shown in Figure 2. Fixations were mainly distributed in four AOIs, and its distribution proportion was larger in AOI3 of interface 1 than AOI3 of interface 2.

RMANOVA by SPSS 13.0 showed that on the 0.05 significant level, the main effect of the information amount was significant for accuracy ($P<0.001$), subjective score ($P<0.001$), and fixation time ($P=0.022$), but insignificant for blink frequency ($P=0.130$). The time factor significantly affected accuracy ($P<0.001$), subjective score ($P<0.001$), blink frequency ($P=0.001$), and fixation time ($P=0.012$). However, accuracy ($P=0.555$), subjective score ($P=0.407$), and fixation time ($P=0.663$) were insignificantly modulated by the display interface, and blink frequency ($P=0.079$) was critical insignificantly modulated by the display interface.

Pearson correlation analysis between the theoretical predictions and the four experimental indices suggested that, under eight MW conditions, predictions were positively correlated with subjective scores ($r^2=0.806$, $P=0.002$) and fixation time ($r^2=0.632$, $P=0.018$), and negatively correlated with accuracy ($r^2=0.778$, $P=0.004$) and blink frequency ($r^2=0.566$, $P=0.032$). On right side of Figure 2,

Table 1
Modeling predictions and experimental results under different MW conditions (mean \pm SD)

MW condition	Prediction/bits/s	Accuracy/%	Subjective Score/point	Blink Frequency/Hz	Fixation Time/s
1	4.734	75.2 \pm 6.8	66.763 \pm 10.533	0.199 \pm 0.111	406 \pm 98
2	2.645	81.8 \pm 7.6	45.738 \pm 13.890	0.268 \pm 0.144	360 \pm 138
3	4.108	79.8 \pm 6.7	51.288 \pm 12.990	0.237 \pm 0.135	374 \pm 130
4	2.296	86.8 \pm 4.1	37.531 \pm 14.844	0.268 \pm 0.152	318 \pm 138
5	3.831	79.9 \pm 8.6	56.469 \pm 8.857	0.311 \pm 0.163	357 \pm 87
6	2.141	82.9 \pm 5.8	46.929 \pm 12.220	0.360 \pm 0.206	354 \pm 84
7	3.302	81.6 \pm 6.2	48.229 \pm 11.656	0.315 \pm 0.174	347 \pm 85
8	1.845	83.9 \pm 7.0	38.784 \pm 16.335	0.381 \pm 0.209	341 \pm 75

Note: MW conditions are 1(1,1,1), 2(1,1,2), 3(1,2,1), 4(1,2,2), 5(2,1,1), 6(2,1,2), 7(2,2,1), and 8(2,2,2), where parenthesis means (display interface, information amount, time).

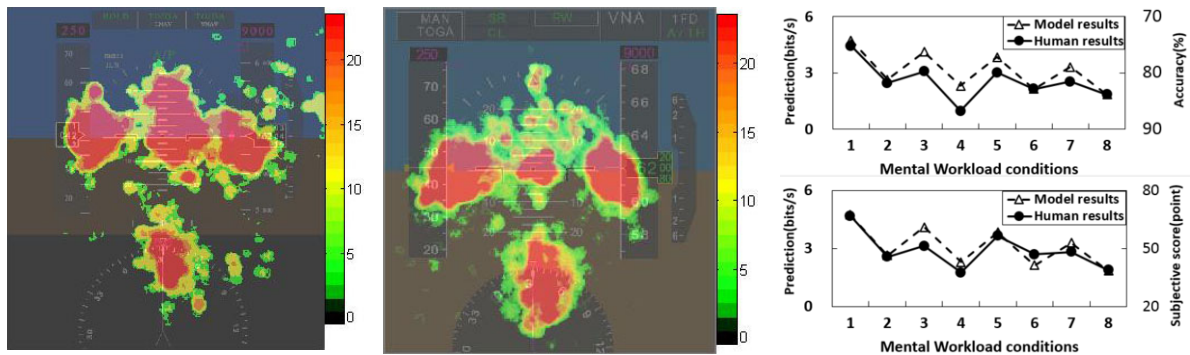


Fig. 2. Attentional fixation heat map (left) and trend of modeling and experimental results (right).

trends of accuracy and subjective score were compared with the trend of predictions.

5. Discussion and conclusion

The significance and trend of experimental subjective scores, performance, and eye movement tracking indicate that the experimental measure indices have certain sensitivity for MW measurement. Both the information amount and time pressure have a significant effect on MW. NASA-TLX is a multi-dimensional MW evaluation scale developed by NASA. Because its validity matches the performance, NASA-TLX is widely harnessed in measuring and diagnosing MW [16,17]; the information cognitive performance is one of the intuitive indices to estimate display interface design. Ideal design should maintain efficient cognition under different information amount conditions. The accuracy of the abnormal attitude recovery task can be taken as a symbol of information cognitive performance [18]. The physiological indices' blink frequency and fixation time are sensitive to visual monitoring workload. Related studies found that a higher operator MW will give rise to a lower blink frequency and more fixation time [17-19]. The absence of the display interface effect on the experimental indices in the present study may be because interface 1 simulates the B787 PFD interface that has already been put into use, is widely recognized, and meets the airworthiness requirements. Compared to the optimized interface 2, interface 1 will not significantly increase the pilot's MW.

The correlation between the MW theoretical prediction and the experimental results suggests that a large amount of information, high time pressure, and improper visual coding may result in a significant increase in the operator's MW, which verifies the validation of mental workload prediction based on attentional resource allocation and information processing. The predictions are highly correlated with subjective score and performance ($r^2=0.81/0.78$), but not highly correlated with the eye movement tracking results ($r^2=0.63/0.57$). One reasonable explanation of these results is that physiological measurements are less sensitive than subjective evaluation and performance. The other possible explanation is that the initial setting of the eye tracker excludes the data with blink durations less than 10ms. Otherwise, the difference between the visual coding designs of two display interfaces is relatively small; hence, the prediction is at a higher level under different information amounts and time pressure conditions than under the two display interfaces. In conclusion, at the early stage of aircraft cockpit display interface design, mental workload prediction based on attentional resource allocation and information processing can predict the pilot's MW when monitoring the display interface and conducting human-rating analysis of cockpit display interface and flight tasks.

In contrast to most existing mental workload models, the mental workload prediction based on attentional resource allocation and information processing raised in the present study adds the attentional allocation factor and considers the relationship between attentional resource allocation and MW; this all provides the model with better prediction precision. Moreover, the model combines the attentional resource demand factor and information processing factor and is quantified by exogenous task demands and endogenous resource supply, which can account for the lack of interference from time-consuming but minimally resource-demanding tasks. Therefore, this model has an expanded scope of application compared to IP model. So far, since the model has not taken advantage of the conflicts within task types and multiple tasks, subsequent research will focus on optimizing the mental workload prediction model.

Acknowledgment

This study is supported by National Natural Science Foundation of China (Program Grant No. 71301005), Research Fund for the Doctoral Program of Higher Education of China (20121102120013) and Aviation Science Foundation of China (20148041003).

References

- [1] M.A. Vidulich and P.S. Tsang, Mental workload and situation awareness, in: *Handbook of human factors and ergonomics*, G. Salvendy, ed., Wiley Press, USA, 2012, pp. 243-273.
- [2] R. Parasuraman, T.B. Sheridan and C.D. Wickens, Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs, *Journal of Cognitive Engineering and Decision Making* **2** (2008), 140-160.
- [3] K.R. Laughery, J.B. Plott, M. Matessa and S. Archer, Modeling human performance in complex systems, in: *Handbook of human factors and ergonomics*, G. Salvendy, ed., Wiley Press, USA, 2012, pp. 931-961.
- [4] R. Parasuraman and E. Rovira, Workload modeling and workload management: Recent theoretical developments, *George Mason University Fairfax VA Department of Psychology* **1** (2005), 1-19.
- [5] C.D. Wickens, Multiple resources and mental workload, *Human Factors: The Journal of the Human Factors and Ergonomics Society* **50** (2008), 449-455.
- [6] J.M. Colombi, M.E. Miller, M. Schneider, et al, Predictive mental workload modeling for semiautonomous system design: implications for systems of systems, *Systems Engineering* **15** (2012), 448-460.
- [7] S.F.M. Liang, C.L. Rau, P.F. Tsai, et al, Validation of a task demand measure for predicting mental workloads of physical therapists, *International Journal of Industrial Ergonomics* **44** (2014), 747-752.
- [8] B. Xie and G. Salvendy, Prediction of mental workload in single and multiple tasks environments, *International Journal of Cognitive Ergonomics* **4** (2000), 213-242.
- [9] S. Cao and Y. Liu, Mental workload modeling in an integrated cognitive architecture, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, SAGE Publications **55** (2011), 2083-2087.
- [10] S. Jo, R. Myung and D. Yoon, Quantitative prediction of mental workload with the ACT-R cognitive architecture, *International Journal of Industrial Ergonomics* **42** (2012), 359-370.
- [11] P.W. Crane, Y. Zhou, Y. Sun, et al, Entropy: A conceptual approach to measuring situation-level workload within emergency care and its relationship to emergency department crowding, *The Journal of Emergency Medicine* **46** (2014), 551-559.
- [12] J.L. Szalma and G.W.L. Teo, Spatial and temporal task characteristics as stress: A test of the dynamic adaptability theory of stress, workload, and performance, *Acta Psychologica* **139** (2012), 471-485.
- [13] L. Zhang, D.M. Zhuang and X.R. Wanyan, Information coding for cockpit human-machine interface, *Chinese Journal of Mechanical Engineering* **24** (2011), 707-712.
- [14] X. Xiao, X.R. Wanyan and D.M. Zhuang, Comprehensive evaluation model of multidimensional visual coding on display interface, *Journal of Beijing University of Aeronautics and Astronautics* (2015). (in press)

- [15] X.R. Wanyan, D.M. Zhuang, H.Y. Wei and J.H. Song, Pilot attention allocation model based on fuzzy theory, *Computers & Mathematics with Applications* **62** (2011), 2727–2735.
- [16] M. Mouz -Amady, E. Raufaste, H. Prade, et al, Fuzzy-TLX: Using fuzzy integrals for evaluating human mental workload with NASA-Task Load indeX in laboratory and field studies, *Ergonomics* **56** (2013), 752-763.
- [17] Z.M. Wei, D.M. Zhuang, X.R. Wanyan, et al, A model for discrimination and prediction of mental workload of aircraft cockpit display interface, *Chinese Journal of Aeronautics* **27** (2014), 1070-1077.
- [18] G. Borghini, L. Astolfi, G. Vecchiato, et al, Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness, *Neuroscience & Biobehavioral Reviews* **44** (2014), 58-75.
- [19] X.R. Wanyan, D.M. Zhuang and H. Zhang, Improving pilot mental workload evaluation with combined measures, *Bio-medical Materials and Engineering* **24** (2014), 2283-2290.