Nuclear Engineering and Technology 51 (2019) 453–462

Contents lists available at ScienceDirect

Nuclear Engineering and Technology

journal homepage: www.elsevier.com/locate/net

Original Article

Comprehensive evaluation method for user interface design in nuclear power plant based on mental workload

Yu Chen [a], Shengyuan Yan [b], Cong Chi Tran [b, c, *]

[a] Heilongjiang University of Science & Technology, Harbin, China
[b] Harbin Engineering University, Harbin, China
[c] Vietnam National University of Forestry, Hanoi, Viet Nam

ARTICLE INFO

Article history:
Received 19 June 2018
Received in revised form
10 September 2018
Accepted 9 October 2018
Available online 10 October 2018

Keywords:
Mental workload
Interface design
Comprehensive evaluation method
Fuzzy comprehensive evaluation
Nuclear power plant

ABSTRACT

Mental workload (MWL) is a major consideration for the user interface design in nuclear power plants (NPPs). However, each MWL evaluation method has its advantages and limitations, thus the evaluation and control methods based on multi-index methods are needed. In this study, fuzzy comprehensive evaluation (FCE) theory was adopted for assessment of interface designs in NPP based on operators' MWL. An evaluation index system and membership functions were established, and the weights were given using the combination of the variation coefficient and the entropy method. The results showed that multi-index methods such as performance measures (speed of task and error rate), subjective rating (NASA-TLX) and physiological measure (eye response) can be successfully integrated in FCE for user interface design assessment. The FCE method has a correlation coefficient compared with most of the original evaluation indices. Thus, this method might be applied for developing the tool to quickly and accurately assess the different display interfaces when considering the aspect of the operators' MWL.

1. Introduction

The human factor is one of the key factors in the design and evaluation of the safety of nuclear power plant (NPP) systems [1]. Numerous statistics show that deficiencies in human factors such as poor control room design, procedure, and training, are significant contributing factors to NPPs incidents and accidents [2]. Hitherto, the number of occurred nuclear incidents is relatively small, but if the serious accident happens, it will be a big disaster for people and environment [3]. Therefore, improving the NPPs safety based on the understanding of human as well as technology characteristics is a permanent research topic.

Operating in NPPs is often a high MWL task. To avoid incidents, the operators must constantly acquire and process much information from their eyes, ears, and other sensory organs. Such information includes detect abnormalities, and respond quickly to events occurring. Especially during emergency situations, operators often need to process large amounts of information in a short time and quickly make the decisions, which easily lead to generation of very high cognitive load, and effect on the efficiency and reliability of the whole system. Presently, main control rooms of NPPs are developing to adopt computer-based human–system interfaces (HSIs). The changes in display and control have affected all aspects of NPP operating functions. Therefore, good interface design will enable operators to accomplish their tasks efficiently and effectively with minimal errors [4,5].

MWL has become a crucial issue for many kinds of the industry since the 1960s. It has been used as an essential factor in the evaluation of human performance in complex systems, and Nachreiner [6] concluded that human performance could be affected by too high or too low MWL. Moray [7] also determined that optimizing the MWL allocation could reduce human errors, increase system safety, and improve operators' satisfaction. Many studies have been carried out on the measurement and assessment of MWL on human-computer interface (HCI) of NPPs [8,9], which mainly employed the subjective evaluation method, performance measurement method, and physiological evaluation method [10]. These studies were mostly based on the results of single-index

* Corresponding author. College of Mechanical and Electrical Engineering, Harbin Engineering University, NO.145-1, Nantong Street, Nangang District, Harbin, 150001, China.
E-mail addresses: trancongchi_bk@yahoo.com, trancongchi@hrbeu.edu.cn (C.C. Tran).

https://doi.org/10.1016/j.net.2018.10.010

measurement and applied statistical methods for analysis and evaluation, and they indicated that different evaluation methods should be applied to different task situations and workload levels. However, no method works in all cases because each method has its advantages and limitations. Therefore, using multi-index to comprehensively evaluate MWL is a reasonable method as an alternative to the single method or index-based evaluation.

In the stage of designing or improving HCI for NPPs, the quick, accurate and comprehensive evaluation of MWL of operator under different display interfaces would play an essential role in optimizing the mental task design of HCI. It is of great importance for decision makers to correctly understand the results of the assessment. For example, when they want to make the decision for the HCI designs, decision-makers have to consider what information they need to make right choices, and it will be difficult for people who are not acquainted with professional knowledge. However, in fact, decision-makers are not required to understand these details in many cases completely. Because when selecting the HCI designs, they probably need a judgment of "good" or "bad," rather than crisp comparisons. Therefore, the systematic evaluation method based on the numerical model and mathematical methodologies need to widely apply in research and comprehensive evaluation.

The comprehensive evaluation is considered as a multi-criteria decision-making issue and has some numerical models applied in this aspect such as the artiflcial neural network [11–13], principal component analysis [14]. In these methods, the Fuzzy theory can be better-managed vagueness or information full of uncertainties [15,16], especially in the management and application phase after evaluation. Fuzzy comprehensive evaluation (FCE) is a method based on the theory of fuzzy sets that was introduced in the 1960s [17]. It has now become an effective multi-factor decision-making tool for comprehensive evaluations. This method has been widely applied in the research and application of multi-criteria comprehensive evaluation in many fields, including, but not limited to environmental [18], petroleum project [19], water quality [20], manufacturing [21,22], safety engineering [23,24], and human-machine interface design [25]. These studies concluded that the sensitivity of FCE method is higher compared with other evaluation methods thanks to the predetermined weights and decreased fuzziness by establishing membership functions. However, FCE method has not yet been studied for its application in MWL assessment and this has been identified as a research gap.

The MWL measurement methods can be divided into three main categories such as performance measures, physiological measures and subjective ratings. However, there is no universal method because each method has its disadvantage. Additionally, the human MWL is influenced by many factors such as task demand, physiology, or working environment. These factors have special characteristics, such as nonlinear, interactive, and fuzzy correlative, that are suitable for multi-factor assessment using the FCE method. Therefore, the objective of this study was to develop a comprehensive method to assess user interface design in NPPs based on operators' MWL. Firstly, the evaluation index system (EIS) was selected based on the MWL measurement methods, including the performance measures (obtained through the speed of task and error rate), subjective rating (NASA-TLX) and physiological measure (eye response). The weight vector of the EIS was obtained based on the combination of the variation coefficient and the entropy weight before establishing the FCE model. Then, comprehensive evaluation scores for two interfaces were carried out. The results in this study might be applied for developing the tool to quickly and accurately assess the different display interfaces based on operator's MWL.

2. Fuzzy comprehensive evaluation method

2.1. Evaluation index system (EIS)

The first step of the FCE method is to ascertain the EIS. In this study, the evaluation indices were selected follows three major types of MWL measurement. Performance may be roughly defined as the effectiveness in accomplishing a particular task [26]. The two main methods to measure workload by performance are primary and secondary measures. The basis for using primary and secondary tasks to measure workload is based on the assumption that people have limited resources [27]. The performance measures can be classified into many categories such as accuracy, task time, worst-case performance, domain-specific measures, etc [28]. Among them, speed and error rate of the task are two critical measures of human performance. These indices have been successfully applied to measure MWL in domains including driving [29] and operating in NPP [8,30]. In this study, the speed and error rate while performing the task was calculated because of the following reasons: (1) two indices are easy to collect using video playback feature. (2) speed of task is an important criterion of operating emergency procedures in NPPs. Operators need to process information and to respond with the correct decision quickly. Under the pressure of time, the operator is at risk of overloading the MWL. (3) error of task involve risky behaviors that need to be understood in order to prevent incidents, because that directly relates to the effectiveness and safety of operating of NPPs system.

MWL assessment based on subjective ratings, such as the NASA-Task Load Index (NASA-TLX), subjective workload assessment technique (SWAT), workload profile method, etc., are becoming increasingly important as evaluation tools and have been widely used to assess the workload of the operator. Subjective rating methods are easy to implement, inexpensive and easily administered. However, the limitation of the subjective results is that they can be affected by characteristics of respondents, like biasedness, response sets, mistakes and protest attitudes [31]. de Waard et al. [32] also suggested that subjective ratings alone cannot capture MWL. Hence, use of subjective methods is not recommended when there is a likelihood that the individuals may fake the results, guess randomly, have low motivation or prejudices. In this study, NASA-TLX method was used to evaluate the operators' subjective workload [33]. NASA-TLX is the most widely used methods and has achieved some solid goals in human factors research [34]. This method is a multi-dimensional rating scale that uses six dimensions of workload to provide the diagnostic information about the nature and relative contribution of each dimension in influencing the overall operator workload.

Physiological measures methods (e.g., psychophysiological signals) can be influenced by emotional factors and physical due to apparatuses are attached to user body that constrains user movements and breaks naturalness of interaction. This equipment is also expensive and depending on the capabilities of the purchased system, requiring technical skill and operational experience [35]. However, these methods require a relatively small number of samples and can provide more accurate reports than subjective measures due to using specialized equipment. They can also provide "real-time" assessment results, thus allowing us to quickly and accurately identify usability problems as they occur [36]. The physiological measures can be divided further into central nervous system measures and peripheral nervous system measures [8]. The central nervous system measures (i.e. electrocardiogram) have high reliability in measurement of MWL [37]. However, the applicability of these measures is limited due to the expensive instruments, and the difficulty in mounting the instrument onto subjects. Therefore, the central nervous system measures were not adopted in this study. Peripheral nervous system measures such as eye response as the indices of MWL [38]. Measurement of eye response is one of the objective measurement methods and useful for assessing the operators' MWL and the quality of the design interface in real time [39]. Eye response measure is a technique that captures eye behavior in response to a visual stimulus, suitable for analyzing human behavior [40]. A record of user's eye movements during interaction (i.e., doing a task) with user interface can indicate the locations and how long user looks on the interface. Additionally, about 80% of operators' thoughts can be inferred correctly using his or her eye movement data [41]. Thus eye movements can signal whether the user interface elicits usage problems or facilitates efficient and effective task realization. Additionally, many eye response parameters that are used as MWL measures in operating industrial control.

In this study, four eye response indices that were used as MWL measures include pupil dilation, blink rate, total fixation duration and fixation rate. Eye responses were recorded using the iView X head mounted eye-tracking device (SensoMotoric Instruments, German) with a rate of 50Hz, pupil/corneal reflection $< 0.1°$ and gaze position accuracy$< 0.5°-1.0°$. BeGaze software version 3.0 was used to analyze and process data. Calibration of the eye tracker was performed for every individual at the beginning of the experiment based on the standard five-point approach provided in the iView X tracker. The RCV interface presented on the 17 inches liquid-crystal display monitor, displayed resolution of 1024 × 768 pixels. During the experiment, raw eye tracking data was recorded in the form of XY coordinates according to sampling rate. According to the above analysis, the selected index system for FCE model is illustrated in Fig. 1.

2.2. Normalization of the original evaluating matrix

With n evaluating indicators and m participant count form an original indicators value matrix:

$$X = \begin{bmatrix} x_{11} & x_{12} \dots x_{1n} \\ x_{21} & x_{22} \dots x_{2n} \\ \dots & \dots \quad \dots \\ x_{m1} & x_{m2} \dots x_{mn} \end{bmatrix} \tag{1}$$

Normalization this matrix to get Eq. (2):

$$R = [r_{ij}]_{m\times n} \tag{2}$$

where, r_{ij} is the data of the jth evaluating object on the indicator, and $r_{ij} \in [0,1]$. Among these indicators, to which the bigger the better, there are (3)

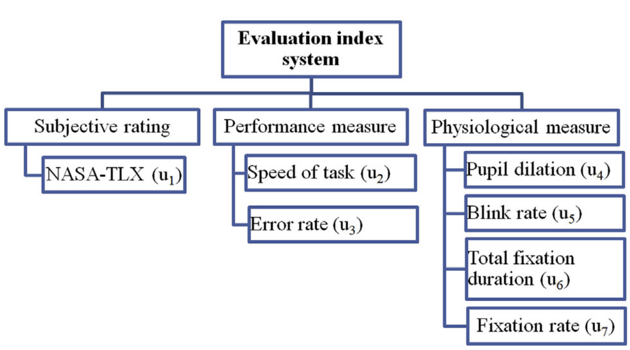

Fig. 1. Evaluation index system based on MWL measurement method.

$$r_{ij} = \frac{x_{ij} - \min_j (x_{ij})}{\max_j (x_{ij}) - \min_j (x_{ij})} \tag{3}$$

while, the smaller the better, there are

$$r_{ij} = \frac{\max_j (x_{ij}) - x_{ij}}{\max_j (x_{ij}) - \min_j (x_{ij})} \Bigg\} \tag{4}$$

2.3. Determining the evaluation criteria

The evaluation criteria for assessing the MWL of user interface should be ascertained. For each index, evaluation criteria and ranks are carefully determined. In this study, the evaluation rank set V assume as: $V_r = [v_1, v_2, \dots v_5] =$ (very low, low, normal, high, very high), respectively with the evaluation criteria is the score, set as $H_r = [1, 2, 3, 4, 5]$. “Very high” means that the operator has maximum MWL when performing the tasks on the interface and “very low” indicates that MWL is negligible.

2.4. Determining the fuzzy relationship matrix

Based on the EIS and assessing rules, the fuzzy relationship matrix is ascertained as:

$$F = \left[f_{ij}\right]_{n\times r} = \begin{bmatrix} f_{11} & \cdots & f_{1r} \\ \vdots & \ddots & \vdots \\ f_{n1} & \cdots & f_{nr} \end{bmatrix} \tag{5}$$

where, f_{ij} represents the fuzzy membership of the ith index belonging to the jth rank.

For correlated indices such as blink rate, the evaluation of MWL is better when the index value increases. Then the fuzzy set function can be selected as:

$$f_{ik}(x) = \begin{cases} 1 & x \geq v_{ik} \\ \dfrac{x - v_{i(k+1)}}{v_{ik} - v_{i(k+1)}} & v_{i(k+1)} < x < v_{ik} \quad \cdot if \cdot k = 1 \\ 0 & x \leq v_{i(k+1)} \end{cases} \tag{6}$$

$$f_{ik}(x) = \begin{cases} 0 & x \leq v_{i(k+1)},\ x \geq v_{i(k-1)} \\ \dfrac{v_{i(k-1)} - x}{v_{i(k-1)} - v_{ik}} & v_{ik} < x < v_{i(k-1)} \quad if \cdot k = 2,3,4 \\ \dfrac{x - v_{i(k+1)}}{v_{ik} - v_{i(k+1)}} & v_{i(k+1)} < x < v_{ik} \end{cases} \tag{7}$$

$$f_{ik}(x) = \begin{cases} 1 & x \leq v_{ik} \\ \dfrac{v_{i(k-1)} - x}{v_{i(k-1)} - v_{ik}} & v_{ik} < x < v_{i(k-1)} \quad if \cdot k = 5 \\ 0 & v_{i(k-1)} \leq x \end{cases} \tag{8}$$

where, x is the select value of the ith index; v_{ik} is the kth rank threshold of the ith index; and f_{ik} is the fuzzy membership of the ith index belonging to the kth rank.

When the MWL is negatively correlated with value of index like NASA-TLX scores, the membership f_{ik} can be given as:

$$f_{ik}(x) = \begin{cases} 0 & x \geq v_{i(k+1)} \\ \dfrac{v_{i(k+1)} - x}{v_{i(k+1)} - v_{ik}} & v_{ik} < x < v_{i(k+1)} \quad if \cdot k = 1 \\ 1 & x \leq v_{ik} \end{cases} \tag{9}$$

$$f_{ik}(x) = \begin{cases} 0 & x \leq v_{i(k-1)}, \ x \geq v_{i(k+1)} \\ \dfrac{x - v_{i(k-1)}}{v_{ik} - v_{i(k-1)}} & v_{i(k-1)} < x < v_{ik} \quad if \cdot k = 2,3,4 \\ \dfrac{v_{i(k+1)} - x}{v_{i(k+1)} - v_{ik}} & v_{ik} < x < v_{i(k+1)} \end{cases} \tag{10}$$

$$f_{ik}(x) = \begin{cases} 0 & x \leq v_{i(k-1)} \\ \dfrac{x - v_{i(k-1)}}{v_{ik} - v_{i(k-1)}} & v_{i(k-1)} < x < v_{ik} \quad if \cdot k = 5 \\ 1 & x \geq v_{ik} \end{cases} \tag{11}$$

2.5. Evaluation index weight analysis

The design of weight is one of the important parts of the fuzzy evaluation method, as it would have a profound effect on the evaluation results. In order to overcome the subjective limitation in traditional FEC method, the index weight of this study is obtained based on the combination of the entropy weight and variation coefficient method. The entropy method is often used for assessing weights in the technique for the order of preference by similarity to ideal solution [42]. This method considers adequately the information of values all the monitoring sections provided to balance the relationship among numerous evaluating objects. However, the distribution of the index weights obtained by the entropy weight method may appear the defects of balance [43]. The variation coefficient method can improve the workload, and overcome the adverse effects of abnormal values [44]. In addition, the combination of variation coefficient and entropy weight method have been shown the effectively adjust the problems existing in the entropy weight method [43,45]. The procedure of calculating the index weight by entropy weighing method is described as follows:

The entropy of ith indicator is defined as:

$$H_i = -\frac{1}{lnn}\sum_{j=1}^{n} f_{ij} ln f_{ij} \quad (i = 1,2,..m) \tag{12}$$

where f_{ij} is the specific gravity value for each r_{ij} and $f_{ij} = \frac{r_{ij}}{\sum_{j=1}^{n} r_{ij}}$ and suppose if $f_{ij} = 0$ then $f_{ij} ln f_{ij} = 0$.

The weight of entropy of ith indicator could be defined as:

$$w_i = \frac{1 - H_i}{m - \sum_{i=1}^{m} H_i} \tag{13}$$

in which $0 \leq w_i \leq 1, \ \sum_{i=1}^{m} w_i = 1$.

The full procedure of determining the weight based on variation coefficient is described as following:

Calculate the mean square deviation of the ith influencing factor:

$$\overline{r_i} = \frac{\sum_{j=1}^{n} r_{ij}}{n} \tag{14}$$

$$\sigma_i = \sqrt[2]{\frac{1}{n-1}\sum_{j=1}^{n}\left(r_{ij} - \overline{r_i}\right)^2}, \ (i = 1,2,..m) \tag{15}$$

where, $\overline{r_i}$ stands for the average value of the ith influencing factor and σ_i is the mean square deviation.

Calculate the variation coefficient of ith influencing factor:

$$E_i = \frac{\sigma_i}{\overline{r_i}}, \ (i = 1,2,..m) \tag{16}$$

Normalized the variation coefficient of each influencing factor and the weights can be obtained

$$\delta_i = \frac{E_i}{\sum_{i=1}^{m} E_i} \quad (i = 1,2,\ldots m) \tag{17}$$

where, $0 \leq \delta_i \leq 1, \ \sum_{i=1}^{m} \delta_i = 1$.

The result of the combination of the variation coefficient and the entropy weight as shown in Eq. (18):

$$\omega_i = \lambda w_i + (1 - \lambda)\delta_i \tag{18}$$

where, λ is the preference factor, and $\lambda \in (0,1)$; $A = [\omega_i] = [\omega_1 \ \omega_2 \ldots \omega_m]$.

2.6. Calculating comprehensive evaluation grade

Fuzzy membership of comprehensive evaluation classes can be calculated by multiplying weight vector and fuzzy relationship matrix.

$$B = A^*F = [\omega_1, \ \omega_2, \ldots, \omega_m]\begin{bmatrix} f_{11} & \cdots & f_{1r} \\ \vdots & \ddots & \vdots \\ f_{m1} & \cdots & f_{mr} \end{bmatrix} = [b_1 \ b_2 \ldots b_r] \tag{19}$$

Fuzzy comprehensive evaluation score is:

$$C = B^*H^T \tag{20}$$

3. Experimental validation

3.1. Participants

This experiment has been conducted with participants who have been trained for the specific task that they have been asked to do. Specifically, twenty-two man engineering students who were 22 ± 1.6 (mean ± SD) years old have been invited to participate in this experiment. They were required to have good vision ability, to be right-handed and had no discomfort on the day of the experiment. The experimental process was overseen by professors and experts who are specialized in human factors and NPPs technology. All participants completed and signed an informed consent form approved by the university institutional review board and were compensated with extra credit in extracurricular activities in their course.

3.2. Chemical and volume control interface

This study used the operation procedure of circuit preparation for S-1-RCV-001 of chemical and volume control (RCV), which is

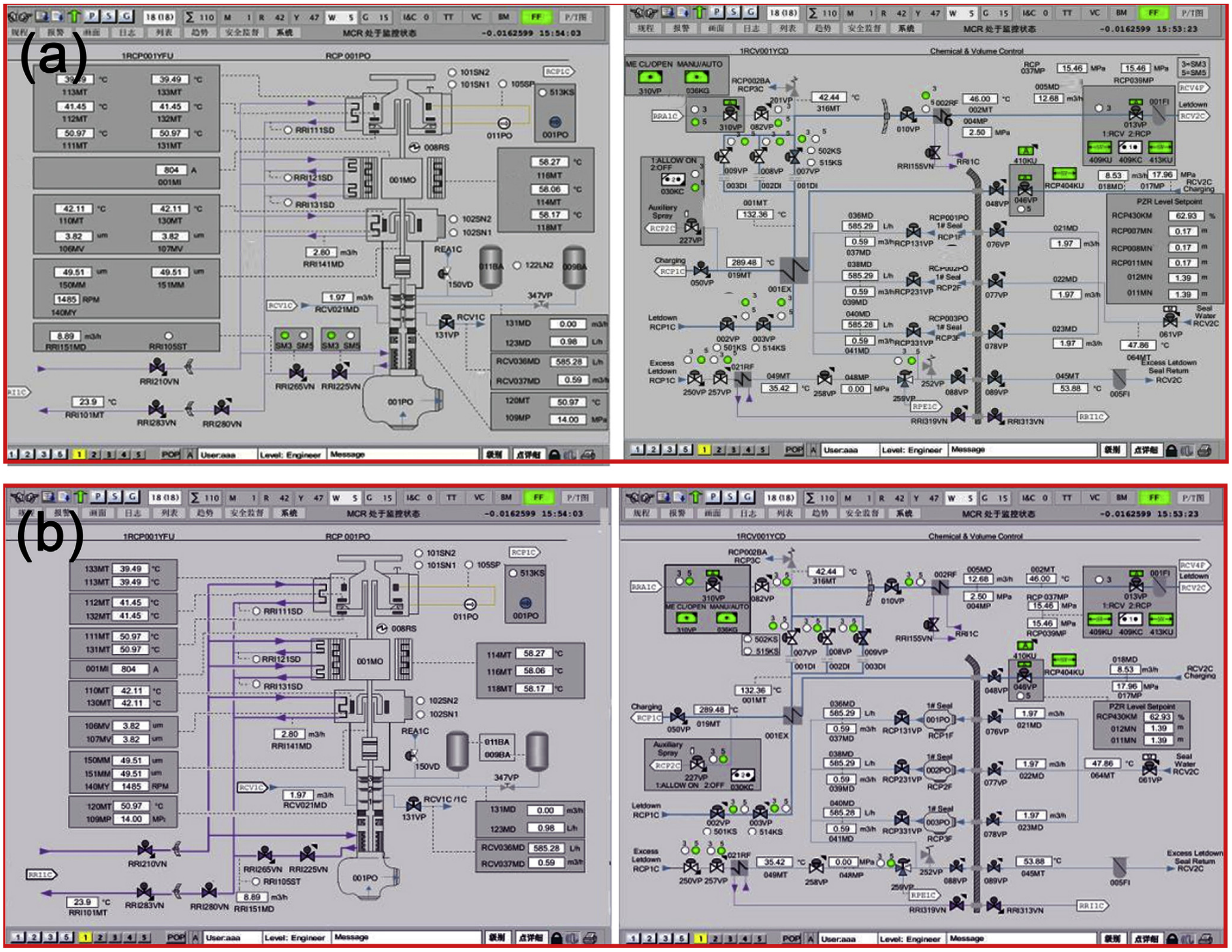

Fig. 2. Original interface (a) and redesigned interface (b) of circuit preparation for S-1-RCV-001.

used in NPP in China (see Fig. 2a). The RCV system helps maintain the primary system water inventory within the acceptable pressurizer level limits in normal operation, during power transients, during start-up of the plant and during heating and cooling transients of the primary. The user interface provides operators the information necessary to monitor the system state. Flow lines connect different system components with arrows that indicate flow direction. Gauge values adjacent to each of valves indicate flow rate, pressure, and temperature. When the interface of RCV system is used, operators must search and locate many valves, pumps place and set required parameters.

The use of human factors and ergonomics approach in the design of process control systems in NPPs presents many opportunities for improvements with regard to system effectiveness, efficiency, reliability and safety [46,47]. Our previous research has improved the limitations of the original RCV interface based on human factors and ergonomics, is presented in Fig. 2b [48]. Although the original RCV design is used in practice, some design characteristics suggest that consideration of avoiding human error and decreasing operators' MWL should be addressed. For instance, the original RCV interface does not specify task content required for completion and it also lacks low-level information, high-level information, and links between reference information support and services. This lack is added based on the criterion of navigational links to related information guide (NUREG-07001.1-27). The criterion of grouping of related information (NUREG-07001.1-28) is used to group the relevant information with the aim to minimize user distraction, which minimizes disadvantages of divided attention and number of attention shifts within a display page and between them. For instance, the data of 120 MT and 109 MT are the same type as the data of 110 MT, 112 MT, 113 MT. However, 120 MT and 109 MT are located on the right side of the main control window, while the indicators that display 110 MT, 112 MT, 113 MT are located on the left side of the screen. It increases the visual distance between objects, making comparisons of the values more difficult. The important information flows in CP001YFU interface are not prominent, namely they are thin and poorly visible. NUREG-07001.2.8-6 (Line Coding) is used to design important display information flows, and NUREG-07001.1-5 (Correspondence Mapping) is used to display information that does not distinguish clearly between other types. In addition, NUREG-07001.1-29 criterion (Spatial Proximity for Related Information) and NUREG-07001.5-7 criterion (Redundancy) are also used to address other issues. The processes of RCV interface analysis and redesigning were conducted carefully and combined with comments from some professors and experts who are specialized in human factors and NPP technology. In addition, user interface layout must also meet the requirements of technology, system functions and operating procedures.

Our previous research has also demonstrated that the redesigned interface is more effective than the original interface in both information searching and processing. The results of *t*-Test analysis indicated that redesigned interface was rated significantly lower in overall workload, time to complete, and fixation duration, but higher blink rate. However, the evaluation results based on statistical analyzes made it difficult for decision-makers, especially for people who are not acquainted with professional knowledge. Thus, this study developed a comprehensive method to assess user interface design. This model might be applied for developing a quick assessment model in interface design field based on operators' MWL.

3.3. Task and procedure

The main objective of this study is to evaluate the interface design based on operators' MWL in NPPs. The participants in this study have been asked to execute one procedure of operation with two S-1-RCV-001 designs, such as original and redesigned interface. The main operations are to be performed, such as search valve/pump, input parameters, and open/close valve/pump, are presented in Table 1 and participants had to perform these actions following the procedure checklist. The process did not have any automatic controllers and had to be monitored and controlled by operators.

Before the experiment, all the participants got about 30 min of training. During the training, the participants read and listened to the electronic lecture, which gave the background information regarding the location and functioning of the system components. Also, they were taught how to use the eye tracking system and how to complete the NASA-TLX questionnaire. The experiment was conducted in a quiet room with fluorescent lighting (neon lights) because the lighting conditions are important in the eye tracking experiment (see Fig. 3). The procedure of the experiment was as follows:

Each participant received the list of operations and practiced to complete the procedure without the eye tracking system and without any time pressure. The practice ended when the participant fully understood the procedure. After that, the participant was asked to answer some questions about the subjective rating methods, eye tracking equipment and operation procedure. The participant was given explanations on anything that he was not clear. Then, the participant took a five minutes rest.

Fig. 3. Experiment environment and equipment.

The participant was provided with the eye-tracker which was validated by the calibration. After calibration step, participant rested his eyes on a blank screen of mild blue colour for 1 min and started the first session on one of two interfaces in a randomized order. In the experiment, participant was allowed to use the list of operations if he needed it.

After finishing the first session, the participant completes the NASA-TLX questionnaire, and then rested for 5 min before he started on the other interface.

The participant repeated steps (1) to (3) using the other interface. Each session lasted approximately 40 min.

All participants were asked to complete task seriously, and no information was given about any specific hypotheses under investigation. The data of the participants who did not complete the tasks were excluded from the statistical analysis.

Table 1
Operation of system boots for S-1-RCV-001.

Step	Task description
1	RCV start
2	Start the RIS012 of PTR water inlet valve
3	Turn off RIS077VP and RIS078VP of booster pump inlet valve
4	Turn off RIS021VP and RIS023VP of hot section injection valve
5	Turn off RIS020VP and RIS029VP of cold section injection valve
6	Turn off RIS032VP and RIS033VP of water tank isolation valve RIS004BA
7	Turn off RIS033VP and 233VP of downstream isolation valve 002BA
8	Turn off 060VP of seal water injection valve
9	Turn off 061VP of seal water control valve
10	Turn on 076VP, 077VP and 078VP valve
11	Turn off 088VP and 089VP
12	Turn off 048VP of safely shell charge pipe valve
13	Start 050VP valve
14	Check the level of 011 MN and 012 MN in control box
15	Run IREA02 or clean the coolant procedure using RCV
16	Start No.1 and turn on 088VP and 089VP valve
17	Maintain the injection flow: 067VP, 068VP and 069VP
18	Use manual control adjust 061VP valve
19	Ensure RRA and PTR piping isolated
20	Increase the reactor coolant pressure using 013VP valve
21	Switch on RCP037MP
22	Change manual control to the automatic mode in 408KU
23	Set 413KU to the INT position, gradually raise to 7 bar
24	Start RCP131VP, RCP231VP and RCP331VP
	…
37	Close the inlet valve RCV034VP in PCV002BA
38	Close the inlet valve RIS013VP in PTR001BA

4. Results

4.1. Data preparation for model development

The evaluation of MWL is becoming increasingly important in designing and using the HCI of many fields. The reason is that when using the interfaces, with the high-level of MWL, human operators may exhibit delayed information processing, or even not can complete the task because the amount of information surpasses their capacity to process it. In contrast, when their MWL is much lower than the proper level, they tend to become bored and easily make mistakes. Therefore, the improvement of HCI design based on operators' MWL is essential, especially in the field of NPPs.

Data on two interfaces, original and redesigned, for the twenty-two participants were collected to prepare for model development and validation, and results are summarized in Table 2. Specifically, the NASA-TLX score was calculated according to three steps. At the first step, weight in each of the six subscales is determined to reveal the priority of the six subscales. All subscales are self-evaluated and selected by the operator in a paired form and 15 different comparisons, and then, each workload dimension is scored 0–5. At the second step, to allocate the rating of workload, each of the six subscales is rated with the goal of determination of each scale's effect on the MWL. At this step, the operator scores each of these six subscales from 1 to 100 based on his/her working condition. At the third and the last step, after determination of weight and rating in

Table 2
The results of all evaluation indices.

Evaluation index	Original interface (Mean ± SD)	Redesigned interface (Mean ± SD)
NASA-TLX score	58.0 ± 7.1	52.3 ± 5.6
Speed (s)	275.6 ± 96.8	231.9 ± 81.4
Error rate (%)	10.2 ± 8.2	9.6 ± 7.1
Pupil dilation (pixel)	53.1 ± 6.9	51.5 ± 5.7
Blink rate (times/min)	18.5 ± 5.2	21.6 ± 6.2
Total fixation duration (s)	87.9 ± 29.7	64.5 ± 21.5
Fixation rate (times/s)	1.57 ± 0.92	1.14 ± 0.64

previous steps, NASA-TLX score is calculated in the range 1–100.

The performance data of participants were also calculated for two interfaces. Specifically, speed of task (sec) was defined as the time that participant spent on the experiment was recorded as complete task time. Error rate (%) was determined by comparing the incorrect sequence of actions to each participant's actual actions. In the eye response measure method, all data of twenty-two participants were processed in order to remove artifacts and gaps such as blinks or eye tracker loses the ability to track when participants moving their heads. Fixation rate was collected within the fixation lengths that varied from 100 ms to 600 ms. The indices with their respective parameters outside the condition have been neglected in further analysis. Additionally, the area of interest (AOI) has been defined as the simulation screen. Thus, all the data fell outside the defined AOI have been excluded from analysis. In this study, pupil dilation unit has been collected by pixel; the average blink rate has been defined as a number of blinks per second. Fixation rate is the number of fixations divided by second.

4.2. Evaluation index weight

Using the data shown in Table 2 and based on Eqs. (12) - (13), the fuzzy relationship matrix was established as Eq. (21):

$$w = [0.126\ \ 0.106\ \ 0.077\ \ 0.127\ \ 0.282\ \ 0.177\ \ 0.105] \tag{21}$$

Based on Eq. (14)– (17), the variation coefficient matrix was established as Eq. (22):

$$\delta = [0.127\ \ 0.132\ \ 0.094\ \ 0.144\ \ 0.224\ \ 0.162\ \ 0.117] \tag{22}$$

Combination weighting for FCE model was established as Eq. (23) based on Eq. (18):

$$\omega = [0.126\ \ 0.119\ \ 0.085\ \ 0.136\ \ 0.253\ \ 0.169\ \ 0.111] \tag{23}$$

The weight vector of the EIS is obtained based on the combination of the variation coefficient and the entropy weights are presented in Table 3.

4.3. Fuzzy comprehensive evaluation

Using the data shown in Table 2 and based on Eqs. (5)–(11), the fuzzy membership matrix F was obtained. For example, with the evaluation data of participant 22nd for the redesigned interface, the NASA-TLX score, speed of task, error rate, pupil dilation, blink rate, total fixation duration and fixation rate were 52.2, 368.1s, 2.8%, 52.4 pixel, 33.4 time/min, 76.0s and 0.4 time/s, respectively. So, the resulted fuzzy membership matrix F is:

$$F = \begin{bmatrix} 0 & 0 & 0.04 & 0.96 & 0 \\ 0.49 & 0.51 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0.98 & 0.02 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{24}$$

Using Eqs. (19) and (23), the fuzzy comprehensive assessment matrix B was calculated as Eq. (25).

$$B = [0.042\ \ 0.044\ \ 0.306\ \ 0.125\ \ 0.483] \tag{25}$$

Using Eq. (20), fuzzy comprehensive evaluation score is 3.96, so the MWL assessment result is “High” level. FCE score of all participants in two interfaces were shown in Fig. 4.

The results indicated that the original interface had 19 participants assess higher than “high” level, while redesigned interface only had 9 one. Compared with five level of MWL, the calculate

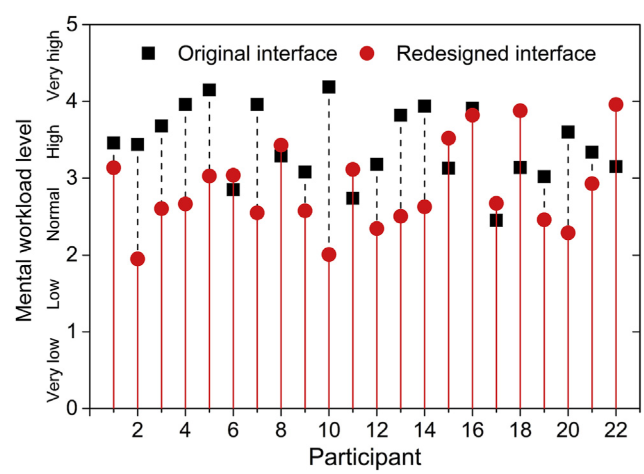

Fig. 4. The FCE score of all participants in two interfaces.

Table 3
The result of combination of weight.

Method	Index	Entropy weight	Coefficient of variation	Combination weighting	
NASA-TLX	Total mental workload	0.126	0.127	0.126	0.126
Performance	Speed (sec)	0.106	0.132	0.119	0.205
	Error rate (%)	0.077	0.094	0.085	
Eye response	Pupil dilation (pixel)	0.127	0.144	0.136	0.669
	Blink rate (times/min)	0.282	0.224	0.253	
	Total fixation duration(sec)	0.177	0.162	0.169	
	Fixation rate (times/sec)	0.105	0.117	0.111	

mean score 3.4 of original interface states that the assessment result of the original interface was relatively high level, reflects that this interface should be improved to reduce the workload for the user. The results also indicated that the score of redesigned interface (2.8) was reduced for 17.6% in comparison to the original interface, which indicates that redesigned interface design significantly reduced the operator's MWL. In addition, the results of *t*-Test of FCE data between two interfaces showed that there are significant differences in average FCE score ($t = 3.43$; $p < 0.01$), indicating that this method may be used to assess the user interface designs in NPPs based on MWL.

4.4. Validation of FCE model

Validation of the FCE model for interface design evaluation was conducted by comparisons between the FCE method and other indices based on the correlation coefficient. To accomplish this, the correlation between FCE score and other indices were tested. Table 4 shows the results of a correlation analysis between FCE score and other indices of original interface design. It showed that the FCE score had a positive correlation with the NASA-TLX score ($r = 0.635$, $p < 0.01$), speed of task ($r = 0.481$, $p = 0.024$), pupil dilation ($r = 0.429$, $p = 0.047$) and total fixation duration ($r = 0.424$, $p = 0.049$), a negative correlation with blink rate ($r = -0.576$, $p < 0.01$).

The correlation between FCE score and other indices of redesigned interface is presented in Table 5. The statistics indicated that FCE score is also correlated significantly with the NASA-TLX score ($r = 0.492$, $p = 0.02$), speed of task ($r = 0.615$, $p < 0.01$), and most eye response data.

5. Discussions

This study was to develop a comprehensive method to assess user interface design in NPPs based on operators' MWL. Performance measurement base on speed of the task and error rate are the easy tools to use. They have been used in evaluations of interface designs in many important industries. In this study, the obtained results of weight have shown that the weight of the task speed has a significant effect (0.119) on the FCE score while the error rate is low (0.085). This result can be explained by the fact that human performance should be considered in terms of both MWL and situation awareness. However, error rate is often not sensitive to workload in a short experimental time [30,49].

The weight value of overall eye response indices was the highest value 0.669, indicating that this measurement method is of particular importance for the evaluation of operator's workload when using the interface control. One of the reasons is that four indices of eye movement in this study directly relate to MWL, and are commonly used in the assessment of MWL. Specifically, blink rate has been observed to decline with increased workload resulting from processing visual stimuli, however, it has been observed to increase with increased load resulting from memory tasks [38] and the connection between blink rate and workload seems tenuous [50]. In this study, blink rate is the highest value (0.253) in eye response indices, which confirms that the redesigned interface is effective in reducing the operators' workload. Pupil dilation indices showed the value 0.136, which is due to pupil dilation can be used as an indirect measurement of the psychological load. It is related to the amount of cognitive control, attention, and cognitive processing required for a given task [51,52]. Recently, a lot of research has used pupil dilation as a factor to consider the amount of MWL in interactive tasks with industrial control interfaces [8,48]. They indicated that pupil diameter in high complexity task is larger than in a low complexity task operation. Fixation indices also can be used to distinguish between workload conditions. Among fixation indices, fixation rate is mostly used by reading, human factors and usability testing. This index is found that it is negatively correlated to difficulty of task, and it can be used as the MWL measurement [53]. Fixation indices in this study have indicated that redesigned interface is more effective in information searching. This was in accord with previous findings that higher fixation on a particular area can be indicative of greater interest in the target, or it can be a sign that the target is complex in some way and more difficult to be searched [54]. This result is also consistent with several recent studies used fixation analysis in MWL evaluation and these indices often show high reliability [55,56]. However, one different challenge in eye movement method is that the data can be influenced by physical, emotional factors and lighting condition. Therefore, this method should be paired with an additional data collection method to interpret the recordings accurately, and subjective assessments are usually used. NASA-TLX has been successfully applied to measure MWL of HCI in many domains such as automobile [57,58], NPPs [8,9,59] and others. The weight value of 0.166 indicated that NASA-TLX method significantly influences the result of the comprehensive evaluation.

The FCE method showed the overall score and level of MWL. The results showed that the operator's MWL level is lower than

Table 4
Correlation between the FCE score and other indices data in original interface.

		FCE score	NASA-TLX	Speed	Error rate	PP	BR	TFD	FR
FCE score	Pearson Correlation	1							
	Sig. (2-tailed)								
NASA-TLX	Pearson Correlation	.635[a]	1						
	Sig. (2-tailed)	.002							
Speed	Pearson Correlation	.481[b]	.361	1					
	Sig. (2-tailed)	.024	.099						
Error rate	Pearson Correlation	.191	.166	-.095	1				
	Sig. (2-tailed)	.395	.460	.675					
PP	Pearson Correlation	.429[b]	.459[b]	-.028	.146	1			
	Sig. (2-tailed)	.047	.032	.903	.516				
BR	Pearson Correlation	-.576[a]	-.414[b]	-.543[a]	.028	-.209	1		
	Sig. (2-tailed)	.005	.043	.009	.902	.351			
TFD	Pearson Correlation	.424[b]	.186	.077	.069	.239	-.476[b]	1	
	Sig. (2-tailed)	.049	.408	.732	.761	.285	.025		
FR	Pearson Correlation	.410	.276	.096	.447[b]	.091	-.010	.180	1
	Sig. (2-tailed)	.058	.214	.671	.037	.688	.965	.423	

PP, pupil dilation; BR, blink rate; TFD, total fixation duration; FR, fixation rate.
[a] Correlation is significant at the 0.01 level (2-tailed).
[b] Correlation is significant at the 0.05 level (2-tailed).

Table 5
Correlation between the FCE score and other indices data in redesigned interface.

		FCE score	NASA-TLX	Speed	Error rate	PP	BR	TFD	FR
FCE score	Pearson Correlation	1							
	Sig. (2-tailed)								
NASA-TLX	Pearson Correlation	.492[b]	1						
	Sig. (2-tailed)	.020							
Speed	Pearson Correlation	.615[a]	.445[b]	1					
	Sig. (2-tailed)	.002	.038						
Error rate	Pearson Correlation	.142	.310	.375	1				
	Sig. (2-tailed)	.528	.160	.085					
PP	Pearson Correlation	.385	.305	.143	.136	1			
	Sig. (2-tailed)	.077	.060	.525	.545				
BR	Pearson Correlation	-.439[b]	-.394[b]	-.228	.197	-.174	1		
	Sig. (2-tailed)	.041	.046	.308	.379	.440			
TFD	Pearson Correlation	.519[b]	.317	.041	-.298	.106	-.211	1	
	Sig. (2-tailed)	.013	.051	.855	.178	.640	.345		
FR	Pearson Correlation	.425[b]	.348	.210	.153	-.046	-.029	.027	1
	Sig. (2-tailed)	.048	.151	.347	.497	.837	.898	.903	

PP, pupil dilation; BR, blink rate; TFD, total fixation duration; FR, fixation rate.
[a] Correlation is significant at the 0.01 level (2-tailed).
[b] Correlation is significant at the 0.05 level (2-tailed).

the original interface when using the redesigned interface, suggesting that the redesigned interface has increased the operator performance. The validity of the proposed method has been tested by correlation between FCE score and other indices. Results analysis showed the correlation between FCE score and most indices of evaluation index system, indicating that FCE method can be used for the interface design evaluation based on MWL and provide an applicable model for MWL control in improving the control interface designs in NPPs. However, the analysis did not reveal a high correlation, possibly due to the influence of some indices less sensitive to workload in this experiment, as well as the limitations of subjective and objective methods when combining them into a single index.

The proposed methodology in this study has some limitations to be mentioned. First, a small number of samples was used in experiments, this reduced statistical power. Additionally, in some cases, the basic visual capabilities may be alike regardless of being an engineering student or a professional operator, all participants were students, who cannot represent the real operators. Although they have basic knowledge, after practice, they had adequate knowledge and skills for solving the problem used in the experiment. Furthermore, their experience might be very different from the experience of real operators who have been trained for years. Finally, experiment environment may cause significant differences among the outcomes; thus results have limited reliability. Therefore, a field study is needed in future to validate the results of the present study because of lighting conditions are essential in eye responses. Also, further validation with professional operators or experts is needed before applying these results in the industry.

6. Conclusions

A fuzzy comprehensive evaluation model was developed to assess user interface design in nuclear power plant (NPP) based on operators' MWL. By using various MWL evaluation methods for reference to establish the index system, the scientific and rationality of the evaluation index system have improved, and the combination of the variation coefficient and the entropy method used to determinate indexes weight can avoid the deviation caused by subjective factor effectively. The results of FCE method indicated that the redesigned interface is more effective than the original interface when considering the aspect of the operators' MWL, which means that different interface designs have a significant effect on the operators' MWL. In addition, correlation analysis showed that FCE correlates with most of the original evaluation indices. This means that the FCE method can perform an effective and integrative assessment of user interface designs based on multiple factors. Thus, this method might be applied for developing the tool to quickly and accurately assess the different display interfaces based on operator's MWL.

Acknowledgements

This study was conducted in the Lab of the College of Mechanical Engineering at Harbin Engineering University, Harbin, Heilongjiang, China. The authors would like to thank the reviewers for their valuable remarks and comments. Also, the authors thank the experts and students who helped conduct this research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.net.2018.10.010.

References

[1] J.V. Hugo, D.I. Gertman, A method to select human–system interfaces for nuclear power plants, Nucl. Eng. Technol. 48 (2016) 87–97.
[2] J. O'hara, J. Higgins, J. Persensky, P. Lewis, J. Bongarra, Human factors engineering program review model, in: Report no.ADA488603, Brookhaven National Lab Upton, NY, 2004.
[3] M. Hatch, E. Ron, A. Bouville, L. Zablotska, G. Howe, The Chernobyl disaster: cancer following the accident at the Chernobyl nuclear power plant, Epidemiol. Rev. 27 (2005) 56–66.
[4] L.H. Ikuma, C. Harvey, C.F. Taylor, C. Handal, A guide for assessing control room operator performance using speed and accuracy, perceived workload, situation awareness, and eye tracking, J. Loss Prev. Process. Ind. 32 (2014) 454–465.
[5] A. Anokhin, A. Ivkin, S. Dorokhovich, Application of ecological interface design in nuclear power plant (NPP) operator support system, Nucl. Eng. Technol. 50 (2018) 619–626.
[6] F. Nachreiner, Standards for ergonomics principles relating to the design of work systems and to mental workload, Appl. Ergon. 26 (1995) 259–263.
[7] N. Moray, Mental workload since 1979, Int. Rev. Ergon 2 (1988) 123–150.
[8] Q. Gao, Y. Wang, F. Song, Z. Li, X. Dong, Mental workload measurement for emergency operating procedures in digital nuclear power plants, Ergonomics 56 (2013) 1070–1085.
[9] Y.-T. Jou, T.-C. Yenn, C.J. Lin, C.-W. Yang, C.-C. Chiang, Evaluation of operators' mental workload of human–system interface automation in the advanced nuclear power plants, Nucl. Eng. Des. 239 (2009) 2537–2542.
[10] P.S. Tsang, M.A. Vidulich, Mental Workload and Situation Awareness, Wiley, New York, NY, 2006.
[11] M. Kankal, Ö. Yüksek, Artificial neural network approach for assessing harbor

tranquility: the case of Trabzon Yacht Harbor, Turkey, Appl. Ocean Res. 38 (2012) 23–31.
[12] X. Dou, Y. Yang, Comprehensive evaluation of machine learning techniques for estimating the responses of carbon fluxes to climatic forces in different terrestrial ecosystems, Atmosphere 9 (2018) 83.
[13] L. Yingying, L. Guodong, G. Qiang, X. Yonghai, Radial basis function neural network based comprehensive evaluation for power quality, in: Power System Technology, 2006. PowerCon 2006. International Conference on, IEEE, 2006, pp. 1–5.
[14] C.E. Balas, M.L. Koç, R. Tür, Artificial neural networks based on principal component analysis, fuzzy systems and fuzzy neural networks for preliminary design of rubble mound breakwaters, Appl. Ocean Res. 32 (2010) 425–433.
[15] A. Ergin, İ.H. Özölçer, F. Şahin, Evaluating coastal scenery using fuzzy logic: application at selected sites in Western Black Sea coastal region of Turkey, Ocean Eng. 37 (2010) 583–591.
[16] H.Y. Wu, K.L. Chen, Z.H. Chen, Q.H. Chen, Y.P. Qiu, J.C. Wu, J.F. Zhang, Evaluation for the ecological quality status of coastal waters in East China Sea using fuzzy integrated assessment method, Mar. Pollut. Bull. 64 (2012) 546–555.
[17] L.A. Zadeh, Information and control, Fuzzy sets 8 (1965) 338–353.
[18] Q. Xie, J.-Q. Ni, Z. Su, Fuzzy comprehensive evaluation of multiple environmental factors for swine building assessment and control, J. Hazard Mater. 340 (2017) 463–471.
[19] C.J. Gao, D.P. Li, Fuzzy comprehensive assessment on environmental impact of petroleum project, in: Applied Mechanics and Materials, Trans Tech Publ, 2013, pp. 734–737.
[20] L. Liu, J. Zhou, X. An, Y. Zhang, L. Yang, Using fuzzy theory and information entropy for water quality assessment in Three Gorges region, China, Expert Syst. Appl. 37 (2010) 2517–2521.
[21] W. Chu, Y. Li, C. Liu, W. Mou, L. Tang, A manufacturing resource allocation method with knowledge-based fuzzy comprehensive evaluation for aircraft structural parts, Int. J. Prod. Res. 52 (2014) 3239–3258.
[22] J. Jiao, H. Ren, S. Sun, Assessment of surface ship environment adaptability in seaways: a fuzzy comprehensive evaluation method, Int. J. Nav. Arch. Ocean 8 (2016) 344–359.
[23] F. Jiang, Q. Zheng, W. Shi, The applied research of fuzzy comprehensive evaluation on talent training mode of safety engineering, Procedia Engineering 43 (2012) 425–430.
[24] J. Kang, J. Zhang, J. Gao, Improving performance evaluation of health, safety and environment management system by combining fuzzy cognitive maps and relative degree analysis, Saf. Sci. 87 (2016) 92–100.
[25] D. Li, Y. Sui-huai, W. Wen-jun, Research on fuzzy comprehensive evaluation of human-machine interface layout of driller control room, in: Computational and Information Sciences (ICCIS), 2013 Fifth International Conference on, IEEE, 2013, pp. 1114–1117.
[26] F.G. Paas, J.J. Van Merriënboer, The efficiency of instructional conditions: an approach to combine mental effort and performance measures, Hum. Factors 35 (1993) 737–743.
[27] Y.-Y. Yeh, C.D. Wickens, Dissociation of performance and subjective measures of workload, Hum. Factors 30 (1988) 111–120.
[28] V.J. Gawron, Human Performance, Workload, and Situational Awareness Measures Handbook, CRC Press, 2008.
[29] S. Yan, C.C. Tran, Y. Wei, J.L. Habiyaremye, Driver's mental workload prediction model based on physiological indices, Int. J. Occup. Saf. Ergon. (2017) 1–9.
[30] S.-L. Hwang, Y.-J. Yau, Y.-T. Lin, J.-H. Chen, T.-H. Huang, T.-C. Yenn, C.-C. Hsu, Predicting work performance in nuclear power plants, Saf. Sci. 46 (2008) 1115–1124.
[31] R.F. Dyer, J.J. Matthews, C.E. Wright, K.L. Yudowitch, Questionnaire construction manual, in: DTIC Document, 1976.
[32] D. de Waard, B. Lewis-Evans, Self-report scales alone cannot capture mental workload, Cogn. Technol. Work 16 (2014) 303–305.
[33] S. Hart, L. Staveland, Development of NASA-TLX (task load index): results of empirical and theoretical research, Hum. Mental Workload (1988) 139–183.
[34] S. Rubio, E. Díaz, J. Martín, J.M. Puente, Evaluation of subjective mental workload: a comparison of SWAT, NASA-TLX, and workload profile methods, Appl. Psychol. 53 (2004) 61–86.
[35] Y. Lean, F. Shan, Brief review on physiological and biochemical evaluations of human mental workload, Hum. Factors Ergon. Manuf. 22 (2012) 177–187.
[36] T.Q. Tran, R.L. Boring, D.D. Dudenhoeffer, B.P. Hallbert, M.D. Keller, T.M. Anderson, Advantages and disadvantages of physiological assessment for next generation control room design, in: Human Factors and Power Plants and HPRCT 13th Annual Meeting, 2007 IEEE 8th, IEEE, 2007, pp. 259–263.
[37] T. Heine, G. Lenis, P. Reichensperger, T. Beran, O. Doessel, B. Deml, Electrocardiographic features for the measurement of drivers' mental workload, Appl. Ergon. 61 (2017) 31–43.
[38] G.F. Wilson, An analysis of mental workload in pilots during flight using multiple psychophysiological measures, Int. J. Aviat. Psychol. 12 (2002) 3–18.
[39] J.L. Rosch, J.J. Vogel-Walcutt, A review of eye-tracking applications as tools for training, Cognit. Technol. Work 15 (2013) 313–327.
[40] D. Toker, C. Conati, B. Steichen, G. Carenini, Individual user characteristics and information visualization: connecting the dots through eye tracking, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2013, pp. 295–304.
[41] J.S. Ha, Y.-J. Byon, J. Baek, P.H. Seong, Method for inference of operators' thoughts from eye movement data in nuclear power plants, Nucl. Eng. Technol. 48 (2016) 129–143.
[42] Z.-H. Zou, Y. Yun, J.-N. Sun, Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment, J. Environ. Sci-china 18 (2006) 1020–1023.
[43] X. Gai, The design and evaluation of ship navigation display and control system based on cognitive load, in: Harbin Engineering University Thesis, China, 2015.
[44] D. Liu, Z. Zou, Water quality evaluation based on improved fuzzy matter-element method, J. Environ. Sci. (China) 24 (2012) 1210–1216.
[45] L.I. Xiao-Xing, D.U. Jun-Kai, F.U. Yao, Water quality evaluation model of normal cloud based on coefficient variation and entropy weight, Water Resour. Power 10 (2017) 55–58. China.
[46] P.V. Carvalho, J.O. Gomes, M.R. Borges, Human centered design for nuclear power plant control room modernization, in: 4th Workshop HCP Human Centered Processes, Citeseer, 2011.
[47] F. Nachreiner, P. Nickel, I. Meyer, Human factors in process control systems: the design of human–machine interfaces, Saf. Sci. 44 (2006) 5–26.
[48] S. Yan, C.C. Tran, Y. Chen, K. Tan, J.L. Habiyaremye, Effect of user interface layout on the operators' mental workload in emergency operating procedures in nuclear power plants, Nucl. Eng. Des. 322 (2017) 266–276.
[49] T.M. Lanzetta, W.N. Dember, J.S. Warm, D.B. Berch, Effects of task type and stimulus heterogeneity on the event rate function in sustained attention, Hum. Factors: J. Hum. Fact. Ergon. Soc. 29 (1987) 625–633.
[50] M. Castor, E. Hanson, E. Svensson, S. Nählinder, P. LeBlaye, I. MacLeod, N. Wright, J. Alfredson, L. Ågren, P. Berggren, GARTEUR Handbook of Mental Workload Measurement, GARTEUR, Group for Aeronautical Research and Technology in Europe, Flight Mechanics Action Group FM, 2003, p. 164. AG13.
[51] S. Moresi, J.J. Adam, J. Rijcken, P.W. Van Gerven, H. Kuipers, J. Jolles, Pupil dilation in response preparation, Int. J. Psychophysiol. 67 (2008) 124–130.
[52] S.M. Wierda, H. van Rijn, N.A. Taatgen, S. Martens, Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution, Proc. Natl. Acad. Sci. Unit. States Am. 109 (2012) 8456–8460.
[53] K.F. Van Orden, T.-P. Jung, S. Makeig, Combined eye activity measures accurately estimate changes in sustained visual task performance, Biol. Psychol. 52 (2000) 221–240.
[54] M. Nakayama, K. Takahashi, Y. Shimizu, The act of task difficulty and eye-movement frequency for the'Oculo-motor indices, in: Proceedings of the 2002 Symposium on Eye Tracking Research & Applications, New Orleans, Louisiana, 2002, pp. 37–42.
[55] U. Ahlstrom, F.J. Friedman-Berg, Using eye movement activity as a correlate of cognitive workload, Int. J. Ind. Ergon. 36 (2006) 623–636.
[56] L.L. Di Stasi, A. Antolí, J.J. Cañas, Evaluating mental workload while interacting with computer-generated artificial environments, Entertain. Comput. 4 (2013) 63–69.
[57] P. Lehrer, M. Karavidas, S.E. Lu, E. Vaschillo, B. Vaschillo, A. Cheng, Cardiac data increase association between self-report and both expert ratings of task load and task performance in flight simulator tasks: an exploratory study, Int. J. Psychophysiol. 76 (2010) 80–87.
[58] A. Pauzié, Evaluation of the driver's mental workload: a necessity in a perspective of in-vehicle system design for road safety improvement, Cognit. Technol. Work 16 (2014) 299–302.
[59] M. Naderpour, J. Lu, G. Zhang, A safety-critical decision support system evaluation using situation awareness and workload measures, Reliab. Eng. Syst. Saf. 150 (2016) 147–159.