International Journal of Technology and Human Interaction
Volume 18 • Issue 1

Development of an Eye Response-Based Mental Workload Evaluation Method:

A Study of User interface in a Nuclear Power Plant

Cong Chi Tran, Vietnam National University of Forestry, Vietnam*
https://orcid.org/0000-0003-1460-4430

Shengyuan Yan, Harbin Engineering University, China
https://orcid.org/0000-0002-0580-2892

ABSTRACT

This study proposed an eye responses-based mental workload (E-MWL) evaluation method in nuclear power plants (NPPs) when performing the task via a user interface control. The fuzzy theory was used to combine four eye response indices using the entropy weight method. Then, the E-MWL method was validated through experiments by comparison with the NASA-TLX rating and performance measures indices in two different tasks of the state-oriented procedure (SOP) in NPP. The correlation analysis results between the NASA-TLX and eye response indices showed that four eye response indices used in this study were correlated significantly with the NASA-TLX, indicating that these indices may develop the E-MWL method. The E-MWL score results indicated that it is highly correlated with NASA-TLX and performance measures indices in two different tasks of SOP in NPP. This has proved that E-MWL is an objective method suitable for evaluating and predicting human mental workload (MWL) for interface control task in NPPs.

KEYWORDS

Eye Response, Human-Computer Interface, Mental Workload, Nuclear Power Plant

INTRODUCTION

The study of human mental workload (MWL) is not new; it has been discussed and researched since 1960s (Kum et al., 2007). It is widely used in the study of human factors and ergonomics for industry fields due to both excessive and low level of MWL could decrease work performance (Nachreiner, 1995). Increasing operators' MWL, or overload, is one of the possible causes of information processing disruptions since the amount of information exceeds their processing capacity. In contrast, a low level of MWL can cause boredom and tend to make mistakes (Ryu & Myung, 2005). With the rapid development of science and technology, sophisticated industrial systems have progressed, in which operators often receive massive MWL task, especially for complex operating procedures in nuclear

DOI: 10.4018/IJTHI.299071
*Corresponding Author

power plants (NPPs) (Hsieh et al., 2015). Although there are many pieces of evidence to show that well-designed information automation can achieve suitable human operator MWL, analyses of various incidents indicate human errors as still a primary cause for more than 70% of accidents in NPPs (Isaac et al., 2002). Thus, enhancing the safety of NPPs based on the level of operators' MWL is an additional significant concern and a permanent research topic.

MWL is also known as "cognitive workload," considers perceptual and cognitive demands in particular, excluding other factors such as physical workload (Hwang et al., 2008). MWL could be defined as "the amount of mental work or effort necessary to perform a task in a given period of time" (Gao et al., 2013; Proctor & Van Zandt, 2018). It is induced not only due to the cognitive demands of the tasks but also by other factors, such as time demands, stress, fatigue and the number and complexity of assigned tasks (Sheridan & Stassen, 1979; Xie & Salvendy, 2000). Many studies have used subjective rating methods, such as the NASA-Task Load Index (NASA-TLX), subjective workload assessment technique (SWAT), workload profile (WP) method, etc., to evaluate the human MWL. The main advantages of subjective rating methods are that results are easy to implement, inexpensive, easily administered and they are provided directly by the operators. However, the disadvantage of this method is that rating results can be affected by characteristics of respondents and context surrounding (Dyer et al., 1976). Furthermore, the subjective workload cannot be collected in real time. Thus, developing the methods for measuring human MWL objectively with directly measured physiological signals is critical and helpful, especially for the industrial control system.

Currently, physiological measure methods are getting more and more attention due to rapid technology development. In contrast with subjective rating methods, physiological methods measure directly over time and can provide more accurate results due to using specialized equipment (Chuang et al., 2016). The basic principle of these methods is based on the response of the body to external sources of workload. They are collected directly and used as physical indices or to consider their correlation with MWL (De Waard, 1996). Some of the common psychophysiological methods to measure MWL are cardiopulmonary activity, eye-based measures, speech activity, brain activity and galvanic skin response. Eye response indices have been used to reflect the temporal distribution workload levels in HCI control task. However, most of the studies focus on the relationship of eye response indices to MWL without any research suggesting a method of combining them into a quantitative workload measurement value. Therefore, the main objective of this study was to develop an eye responses-based mental workload (E-MWL) evaluation method in the task of searching and processing of information in user interface control. The fuzzy comprehensive evaluation and entropy method were proposed to develop the E-MWL method based on the combination of eye response data.

The remainder of this paper is organized as follows: First, we give an overview of eye indices-related measures of human MWL and fuzzy comprehensive evaluation method. Then, we present the process of suggesting the E-MWL method. The next section presents the validation of the suggested measure through experiments. Finally, we discuss the results and closes with a conclusion. This proposed method might be applied to measure a human operator's MWL in the user interface control tasks in NPPs. From this, the manager can organize the human resources for each specific task to maintain suitable MWL as well as to improve the operator's work performance.

RELATED WORK

In the task of monitoring and operating the industrial system via a user interface, eye response measurement is one of the most objective and useful measurement methods for evaluating operators' MWL and the quality of interface designs (Rosch & Vogel-Walcutt, 2013). Various eye response parameters have been used to measure MWL, including Pupil diameter (dilation, size), blink rate, fixation, etc., have previously researched and confirmed as useful estimates of the human MWL. Pupil diameter has often been observed and evaluated in human factors study. It has been found to increase with increasing MWL as well as with a greater degree of difficulty of a task (Batmaz & Ozturk, 2008;

Hampson et al., 2010; Wierda et al., 2012). Iqbal et al. (2004) also concluded that pupil size to be "the most promising single measure of MWL". Recently, many studies have used pupil diameter as a factor to evaluate MWL levels in interactive tasks with user control interfaces. Gao et al. (2013) found that pupil diameter in high complexity emergency operation procedures (EOP) is larger than in a low complexity EOP in NPP operation. Yan et al. (2017) also used this index to evaluate the interface design of NPP based on the searching task in EOP; however, no significant difference in pupil diameter was found in their study. Pupil diameter also applied to many other fields such as aviation, marine, driving, etc (Orlandi & Brooks, 2018; Philippe et al., 2016; Tran et al., 2017). However, although pupil diameter is useful to reflect temporal distribution workload levels in operating task, it is small and easily drowns in the large changes due to variation in light intensity of the surrounding environment or varying brightness of the screen may easily introduce artifacts into the data.

Pupil diameter usually combines with blink indices, fixation, saccadic and dwell time to estimate the human MWL of different tasks (Brookings et al., 1996; Van Orden et al., 2000). The eye blink is believed to be an indicator of both fatigue and workload. Numerous workload studies have considered blink rate (or blink frequency), duration and amplitude. Among them, blink rate has been observed to decline with greater workload due to processing visual stimuli, however, it has been observed to increase with increased load resulting from memory tasks (Wilson et al., 2004) and the connection between blink rate and workload seems tenuous (Castor et al., 2003). Several studies associate the rate of blinks with MWL; however, many studies on the relationship between blink rate and MWL showed the conflicting results. Specifically, there are some researchers failed to find a significant relationship between blink rate and MWL (Casali & Wierwille, 1983; Veltman & Gaillard, 1998). This is due to the results related to blink rate were mixed, with sometimes increasing rates and sometimes decreasing rates depending on the visual demands (Kramer, 1991). However, most studies have found that an increase in blink rate is associated with a decrease in task demands, particularly visual demands of tasks (Gao et al., 2013; Tran et al., 2017; Wilson, 2002; Yan et al., 2017). Hwang et al. (2008) also used eye blink rate and duration as measures of MWL in a simulated reactor shutdown task in NPPs, and they indicated that most of the participants' eye blink duration was shorter and eye blink rate was less during the high task complexity than the low task.

Eye fixations are eye movements that stabilize the retina over a stationary object of interest and calculated by an event-detection algorithm. Fixations are then counted in the selected space, time, or portion of data. Among fixation indices, fixation rate is found that it is correlated to task complexity, and it can be used as the MWL measurement (Di Nocera et al., 2007; Van Orden et al., 2000). Fixation duration is also frequently used to assess MWL. It is concluded to increase with an increase in cognitive workload (Recarte & Nunes, 2000). Recently, several studies used fixation analysis in MWL evaluation and these indices often show high reliability (Ahlstrom & Friedman-Berg, 2006; Di Stasi et al., 2013; Yan et al., 2017). Similar to fixations, saccades have also been used as an indicator relative to MWL. Saccades, defined as ballistic eye movements that occur on very short timescales between fixations (Sibert & Jacob, 2000). Saccades rate increases when task difficulty, MWL or fatigue decreases (Nakayama et al., 2002). This conclusion is confirmed by Pan et al. (2004), who found a difference in saccades rate between two types of web page interface. In general, saccades rate almost identical to the fixation rate.

The fuzzy comprehensive evaluation method is to quantitatively identify the grade of each factor from the factor layer to the target layer. Compared with other methods such as the data-driven artificial neutral network (Kankal & Yüksek, 2012), statistical method (Takuska-Węgrzyn, 2008), support vector machine (Zhou et al., 2012), principal component analysis (Balas et al., 2010) and others, fuzzy theory can better manage vagueness or information full of uncertainties. This method is developed based on the fuzzy set theory introduced in the 1960s (Zadeh, 1965). It is very useful in the management and application stage after evaluation as it is of great importance for decision-makers to correctly understand the evaluation results. Recently, a lot of research has used fuzzy theory as a popular method for comprehensive evaluation method in many fields. Wang et al. (2015) used the fuzzy comprehensive evaluation model to assess operational ocean observing equipment based on 17 typical instruments in waves, water levels and winds in China. Asadzadeh et al. (2013) adopted fuzzy comprehensive model to analyze the integrated health, safety, environmental and ergonomics. The fuzzy theory was selected for the assessment of HCI designs (Chen et al., 2019). Therefore, the fuzzy comprehensive evaluation method has been widely applied in the application of multi-criteria comprehensive evaluation as a scientific quantitative evaluation means.

DEVELOPMENT OF E-MWL METHOD

Suggestion of E-MWL Method

In this study, the fuzzy theory was adopted for the combination of eye response indices using real measurement data, and their weights were given using the entropy method. The theory of fuzzy sets was introduced in the 1960s (Zadeh, 1965). It has now become an effective comprehensive evaluation tool based on multi-factor evaluation. Fuzzy theory has been successfully used in various domains related to environmental assessment (Xie et al., 2017), manufacturing (Chu et al., 2014), safety engineering (Jiang et al., 2012; Kang et al., 2016), engineering design (Jiao et al., 2016)., etc. Almost results of these studies showed that the sensitivity of fuzzy method is high due to pre-determined weights and decreased fuzziness by establishing membership functions (Li et al., 2013).

Human MWL is affected by many factors such as time demands, stress, complexity of tasks, working environment and their physiology. Physiological signal such as eye response were employed in this work as they are generally involuntary and represent objective data points. They have special characteristics, such as nonlinear, interactive, and fuzzy correlations, which are suitable for evaluating multi-factor by fuzzy theory. The detailed steps in the method are as follows:

Step 1: Establish the original variable matrix of eye response data A.

With m eye response indices and n participants form an original data. The matrix A can be expressed as Eq. (1):

$$A = \left[a_{ij}\right]_{n\times m} \tag{1}$$

where a_{ij} is the jth evaluating indices value of participant i.

Step 2: Normalization of the eye response data.

Due to the large differences in the dimensions, sizes, and evaluation standards of the indices, the comparability of these indices is poor. Therefore, these indices need to be normalized to achieve good comparability. Normalization data is determined by Eq. (2):

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

where, r_{ij} is the data of the jth eye response indices value of participant i, and $r_{ij} \in \left[0,1\right]$. Between these indices, to which the bigger value is higher MWL, we get Eq. (3):

$$r_{ij} = \frac{a_{ij} - \min_j \left(a_{ij}\right)}{\max_j (a_{ij}) - \min_j \left(a_{ij}\right)} \tag{3}$$

In contrast, to which the bigger values are lower MWL, we get Eq. (4):

$$r_{ij} = \frac{\max_j \left(a_{ij}\right) - a_{ij}}{\max_j \left(a_{ij}\right) - \min_j \left(a_{ij}\right)\}} \tag{4}$$

Step 3: Establish a level and score of E-MWL.

To establish the score for E-MWL, the MWL level is distinguished as 5 degrees, hence, the evaluating score set can be defined as:

$$S = \left[20 \quad 40 \quad 60 \quad 80 \quad 100\right] = [very\ \ low, low, normal, high, very\ \ high \tag{5}$$

Step 4: Determine a weights vector.

The weight design of evaluation indices is one of the critical parts in the fuzzy evaluation method, as it would directly impact the evaluation results. In this study, the entropy method (Shannon, 1948) was used to calculate the weights of pupil diameter, blink rate, fixation rate and saccade rate. Basically, the entropy method is based on the actual data via the mathematical method to get the index weight. This method will balance the relationship between evaluation indices by similarity to an ideal solution based on considering adequately the information of values all the evaluation indices (Zou et al., 2006). Therefore, it can adjust the existing problems compared to the subjective weighting method. The entropy method has been widely used to calculate the weight of evaluation indices in engineering (Chen & Hao, 2011; Wang & Lee, 2009; Zou et al., 2006).

The eye response indices weights vector is determined as in Eq. (6):

$$w = \left(w_1, w_2, w_3, \ldots w_m\right) \tag{6}$$

where the weight of ith indices could be defined as:

$$w_i = \frac{1 - H_i}{m - \sum_{i=1}^{m} H_i} \tag{7}$$

where H_i is calculated by Eq. (8) and $0 \leq w_i \leq 1, \sum_{i=1}^{m} w_i = 1$:

$$H_i = -\frac{1}{\ln n} \sum_{j=1}^{n} f_{ij} \ln f_{ij} \left(i = 1, \ldots, m\right) \tag{8}$$

in which $f_{ij} = \frac{r_{ij}}{\sum_{j=1}^{n} r_{ij}}$ and suppose if $f_{ij} = 0$ then $f_{ij} \ln f_{ij} = 0$.

Step 5: Establish a fuzzy relationship matrix.

The fuzzy relationship matrix is defined by Eq. (9):

$$F = \left[f_{ij}\right]_{m\times 5} \tag{9}$$

where, f_{ij} represents the fuzzy membership of the jth evaluating indices value of participant i.

The triangular distribution function was selected to develop the fuzzy set function based on eye response evaluation criteria and ranks.

Step 6: E-MWL score can be calculated by Eq. (10):

$$C = \left[B\right]\times\left[S\right]^{T} \tag{10}$$

where, B matrix can be calculated by multiplying weight vector and fuzzy relationship matrix in Eq. (11):

$$B = [w]\times[F] \tag{11}$$

Selection of Eye Response Indices

Eye response data could be selected to develop the quantitative operator MWL measurement for NPPs in the task of searching and processing of information in user interface control for the following reasons: (1) previous studies based on tasks in NPPs reported that the eye response measures are sensitive to variations of MWL in applied settings regarding arousal in visual search performance. (2) many eye response indices show high reliability when applied to MWL assessments in the industrial control interface. (3) currently, eye-tracking devices are becoming more and more popular. In addition, many low-cost eye-tracking devices have also been identified that it is suitable for detecting and measuring changes in the cognitive load of the operator (Čegovnik et al., 2018). Therefore, this study used four eye response indices to develop the E-MWL include pupil diameter, blink rate, fixation rate and saccade rate. Pupil diameter unit has been collected by mm; the average blink rate has been defined as a number of blinks per second. Fixation rate is the number of fixations divided by second, and the saccade rate has been measured as the number of saccades per second.

The evaluation criteria was proposed based on the basis of eye response indices in previous studies (Bentivoglio et al., 1997; Guillon et al., 2016; Nyström et al., 2016; Peshkovskaya et al., 2017; Wang et al., 2011; Wyatt, 1995), as well as suggestions from some experts who are specialized in human factors. The evaluation criteria and ranks show in Table 1.

EXPERIMENTAL VALIDATION OF E-MWL

This experiment was conducted to confirm whether the proposed MWL measurement method based on the combination of eye response indices can evaluate the human MWL and whether it is superior to existing methods. To accomplish this, the correlation between performance measures method and

Table 1. The eye response evaluation criteria and ranks

Indices	Evaluation criteria				
	Very low	Low	Normal	High	Very high
Pupil diameter	2.5	3	3.5	4	4.5
Blink rate	1.25	1.0	0.75	0.5	0.25
Fixation rate	0.50	1.0	1.5	2.0	2.5
Saccade rate	2.0	1.55	1.1	0.65	0.2

subjective rating method with E-MWL were tested in two tasks of operation with different complexity. The statistical analysis was conducted using SPSS software, version 20. In all cases, α level of 0.05 was used to determine statistical significance.

Selection of Measures Method

Performance Measures Method

The human MWL can be evaluated by performance measures, physiological measures, and subjective ratings (Tsang & Vidulich, 2006). Performance can be defined as the effectiveness in accomplishing a particular task. This method involves the collection of data from one or more subjects performing the task or tasks of interest using primary task performance and secondary task performance. In most investigations, the primary task performance will always be of interest as its generalization to the in-service performance is central to the study while the performance of the secondary task itself may have no practical importance. However, primary tasks are not very sensitive to changes of workload, especially when operators have spare capability to increase their effort level (Choi et al., 2018). The measure of secondary task provides a more sensitive measurement of operator capacity compared to the measure of primary task. However, it has the drawback that the measurement itself contaminates human performance by interfering with primary tasks. In addition, it is difficult to find a secondary task that matches primary tasks (Wu & Li, 2013).

In this study, operation time (second) and the number of errors were used as performance measures to validate the E-MWL because of following reasons: (1) operation time and the number of errors of operators data cannot fake the results or guess randomly due to the video playback feature collecting. (2) operation time and the number of errors are two important criterions of NPPs operating procedures. (3) numerous studies have been applied two indices to measure MWL in NPP (Gao et al., 2013; Hwang et al., 2008; Jou et al., 2009; Yan et al., 2017). The number of error was defined as the incorrect operations while performing the task and operation time was defined as the time that participant spend on each task of experiment. They were collected using video playback feature and were decided by experts

Subjective Ratings Method

Although subjective rating methods have some limitations, they are considered the easiest method, most convenient, least time consuming, and the least expensive form of evaluating MWL. In this study, the subjective rating method such as the NASA-TLX was used to evaluate the participants' MWL. This method is the most widely used to measure MWL of HCI in many industrial domains such as automobile (Lehrer et al., 2010; Tran et al., 2017; Yan et al., 2019), NPPs (Jou et al., 2009; Naderpour et al., 2016; Yan et al., 2017) and others. NASA-TLX is calculated in the range from 0 (very low) to 100 (very high) workload based on six dimensions and their weightings. With this method, the participant provides ratings for a task on six dimensions of workload include mental demand, physical demand, temporal demand, own performance, effort and frustration. Each of dimensions is on a scale of 0 to 100 and their weightings are obtained by fifteen pairwise comparisons of the dimensions.

Participants

In this study, the validation of E-MWL method was conducted with participants who were trained for the specific task due to the difficulty of accessing the experts in NPPs. Thirty-two postgraduate engineering students with an age of 25 ± 3.8 years (M ± SD) were invited to participate in our experiment. Regarding the confirmation of the applied extension of the evaluation results in other future studies, the background information of all the participants, such as age, level of education, and authorization for data collection, was recorded as shown in Table 2. In addition, all participants had good vision ability, right-handed and good health on the day of the experiment. The experimental steps were supervised by experienced experts in the field of human factors and NPPs technology.

Equipment

State Oriented Procedure (SOP) System

The SOP was developed by French institutes (France) around 1980 based on the limitations of the Event Oriented Procedures in the traditional NPPs. After the Three Miles Island accident in 1979, the EOP limitations were apparent. Since then, accidental operating procedures have been extensively studied worldwide. The application of a digital SOP has significantly changed the logic of the operator in handling an accident and the information display pattern in the main control room. Namely, when the SOP system is used, operators have to search and utilize information related to the state parameter, system function, equipment, and procedures (Yang, 2010).

The principle of SOP is based on implicit assumptions in order to reflect actual events better. Specifically, it holds that: (1) events that happen can be none of the predefined events, but a combination of different events; (2) events can happen in a different way from what the event analysis predicts because of the impact of other events, device errors, operation errors, and other related factors; and lastly, (3) the procedure of incidental operating has to be a tool for operators can be diagnosed and identified complex events by providing diagnostic criteria for direct measurement on a continuous or iterative basis at a minimum rate. The safety procedures and injection procedures of the existing SOPs can be divided into three categories: digital operation procedures, hard disk operating procedures, and dynamic devices. The digital operation procedures and some of the hard

Table 2. Demographic profile of participants

Characteristics		Number
Gender	Female	8
	Male	24
Age	23 or under	12
	24-25	11
	26 or older	9
Education level	Master student	26
	Ph.D student	5
	Postdoctoral	1
Vision	Normal vision	23
	Wear glasses	9

disk operating procedures are mixed and arranged in the digital protocol interface, and the display information of dynamic devices are arranged in the screen of special procedures.

The SOP procedures included all plant conditions from the normal power operation to the shutdown, namely hot shutdown, intermediate shutdown, cold shutdown, maintenance shutdown, and refueling shutdown. The interface used in this study is the interface of SOP procedure, which is used in NPP in China, Figure 1. The interface displays system, its different components and the information necessary to monitor the system state.

Eye Tracking System

The eye response data of participants were recorded by the iView X head-mounted eye-tracking device (SMI, Germany) at a sampling rate of 50 Hz, pupil/corneal reflection of less than 0.1°, and gaze position accuracy in the range 0.5°-1.0°. Calibration of the eye tracker was performed for each participant at the beginning of the experiment using the five-point method. The fixation rate was collected within the fixation length that varied from 80 ms to 900 ms, a filter depth of 80 ms, and a saccade length of 20 pixels. The area of interest was defined as the screen of the system interface. During the experiment, first, the raw eye-tracking data were recorded, and then, the data were analyzed by BeGaze software (version 3.0.169). The experimental environment and eye-tracking equipment procedure are shown in Figure 2.

Experimental Task

In the experiment, two operating procedures in the SOP system were used to validate the proposed E-MWL method. Due to MWL is expected to increase in proportion to the increase of a task's complexity [49], the participants have been asked to execute two tasks of operation with low- and high-

Figure 1. The interfaces of RCE (a, b) and SI-CHK (b, d) of SOP system in NPPs

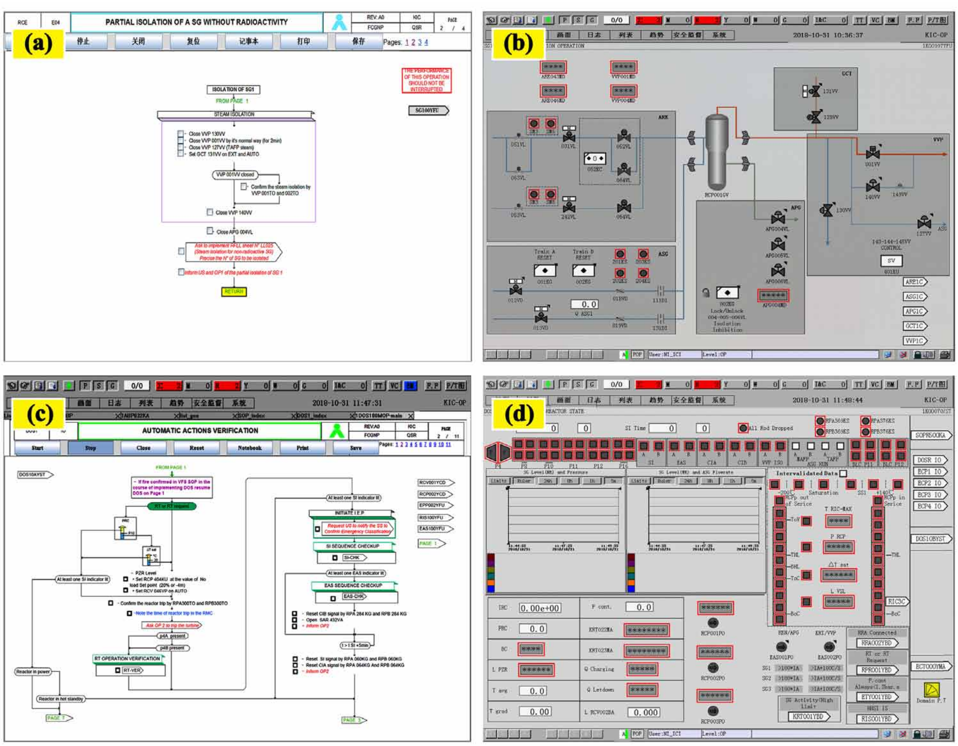
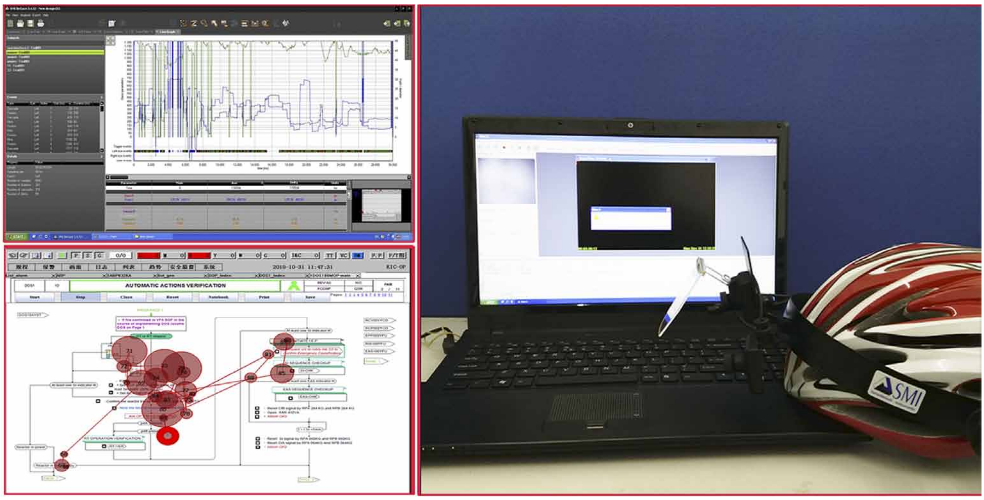


Figure 2. IView X head mounted eye-tracking device and data analysis software



workload levels, such as the partial isolation of a steam generator without radioactivity (abbreviated as RCE) procedure and the safety injection sequence checkup (abbreviated as SI-CHK), respectively. The two procedures differed in the complexity and operation time were proposed by experts who are specialized in NPP technology: the RCE procedure had 28 steps, while the SI-CHK procedure had 21 steps, which are presented in Table 3.

Experimental Procedure

All the participants received about one hour of training on the SOP system before the experiment. They were also taught how to use the iView X head mounted eye-tracking device and how to complete the questionnaire of NASA-TLX method. The experiment was conducted in a quiet room with full light by fluorescent lighting (neon lights). The experiment procedure was as

Table 3. The RCE and SI-CHK operation procedures

RCE operation description		SI-CHK operation description	
1	Close VVP 130VV	1	Confirm SI by RPA 058TO and RPB 058TO
2	Close VVP 001VV by normal way	2	Note the time of SI in the RMC
3	Close VVP 127VV (TAFP steam)	3	Confirm CIA by RPA 062TO and RPB 062TO
4	Set GCT 131VV on EXT and AUTO	4	Confirm RIS 001PO and 002PO IS
5	Confirm the steam isolation by VVP 001TO and 002TO	5	Confirm RIS 077VP and 078VP open
…	….	…	….
25	Close VVP 142VV	18	Confirm RIS 032VP,033VP,034VP and 035VP open
26	Close APG 006VL	19	Confirm RIS 061VP and 062VP open
27	Ask to implement RFLL sheet N° LL025 (Steam isolation for non-radioactive SG)	20	Verify SI memories of diesel reloading by LHA 021KS and LHB 021KS
28	Inform US and OP1 of the partial isolation of SG 3	21	Verify the arming of recirculation memory by RPA 369KS and RPB 369KS

follows. (1) Each participant received the RCE and SI-CHK procedure and practiced to complete them without the iView X head mounted eye-tracking. The practice step ended when participant understood and operated proficiency procedures. To improve the reliability of the results, each participant was interviewed by an expert to ensure that he was ready for the experiment. After that, the participant took a 10 min rest. (2) Each participant was provided eye-tracking device and performed calibration. Then, the participant rested his eyes on a mild blue of the blank screen for 1 min and started the RCE or SI-CHK task in randomized order. (3) After the first task, the participant completed NASA-TLX questionnaire, and then rested for 15 min before starting on task 2. In order to ensure the reliability of evaluation results, all participants were not provided the information about the purpose of the experiment, and they were required to complete the experiment seriously.

Experimental Results

Significant Differences in Two Workload Levels

In this study, Paired t-Test was used to test the difference between NASA-TLX score, performance and eye response indices at two tasks. All the data from the experiment followed approximately normal distribution and there was no significant outliers in the differences between the two related analysis groups. The descriptive statistics are presented in Table 4. At α of 0.05, the result of NASA-TLX score showed that there was a statistically significant difference in task complexity ($t = -4.02$; $p < 0.01$). Performance measures results also showed that average operation time of two task is significantly different ($t = -24.6$; $p < 0.01$). However, the number of error was not significant different ($t = -1.06$; $p = 0.148$). The eye response data of all the participants of the two tasks are also given in Table 3. The results of paired t-Test indicate that there are significant differences in the average pupil diameter ($t = -6.87$, $p < 0.01$), blink rate ($t = 3.69$, $p < 0.01$), fixation rate ($t = -2.37$, $p = 0.012$) and saccadic rate ($t = 3.41$, $p < 0.01$).

Correlation Between the NASA-TLX and Eye Response Indices

Correlation analysis was conducted to consider the relationship between the NASA-TLX score and eye response indices as showed in Tables 5 and 6. The results showed that most eye response indices in this study are correlated significantly with the NASA-TLX. Specifically, the score of NASA-TLX and blink rate were positively correlated in two tasks. The statistics also revealed that the NASA-TLX score is correlated significantly with the saccadic rate in the SI-CHK task. In addition, the higher score of NASA-TLX was also associated with the higher pupil dilation and fixation rate. The results indicated that four eye responses indices used in this study may be used to evaluate MWL.

Table 4. t-Test of NASA-TLX score, performance and eye response indices between two workload levels

Method	Low (RCE procedure) (M ± SD)	High (SI-CHK procedure) (M ± SD)	p-value
NASA-TLX score	50.5±7.9	58.6± 6.4	<0.01
Operation time (s)	202.6±24.2	431.9±46.6	<0.01
Number of errors	2.8±1.7	3.3±1.5	0.148
Pupil diameter (mm)	2.76± 0.32	3.26±0.24	<0.01
Blink rate (blinks/s)	0.45± 0.07	0.35± 0.12	<0.01
Fixation rate (fixation/s)	1.37± 0.58	1.82± 0.86	0.012
Saccade rate (saccade/s)	0.93± 0.44	0.65± 0.20	<0.01

Table 5. Correlation between NASA-TLX and eye response indices in RCE task

		NASA-TLX	Pupil diameter	Blink rate	Fixation rate	Saccade rate
NASA-TLX	Pearson's r	1				
	Sig. (2-tailed)					
Pupil diameter	Pearson's r	.705**	1			
	Sig. (2-tailed)	.000				
Blink rate	Pearson's r	-.602**	-.379*	1		
	Sig. (2-tailed)	.000	.033			
Fixation rate	Pearson's r	.600**	.540**	-.424*	1	
	Sig. (2-tailed)	.000	.001	.016		
Saccade rate	Pearson's r	-.266	-.169	.224	-.407*	1
	Sig. (2-tailed)	.141	.356	.217	.021	

*Correlation is significant at the 0.01 level (2-tailed).
**Correlation is significant at the 0.05 level (2-tailed).

Table 6. Correlation between NASA-TLX and eye response indices in SI-CHK task

		NASA-TLX	Pupil diameter	Blink rate	Fixation rate	Saccade rate
NASA-TLX	Pearson's r	1				
	Sig. (2-tailed)					
Pupil diameter	Pearson's r	.720**	1			
	Sig. (2-tailed)	.000				
Blink rate	Pearson's r	-.868**	-.511**	1		
	Sig. (2-tailed)	.000	.003			
Fixation rate	Pearson's r	.525**	.336	-.433*	1	
	Sig. (2-tailed)	.002	.060	.013		
Saccade rate	Pearson's r	-.593**	-.567**	.613**	-.403*	1
	Sig. (2-tailed)	.000	.001	.000	.022	

*Correlation is significant at the 0.01 level (2-tailed).
**Correlation is significant at the 0.05 level (2-tailed).

E-MWL Evaluation Score

Based on the Eq. (6)-(8), the eye response indices weights of low $\left(W_1\right)$ and high $\left(W_2\right)$ workload levels were established as Eq. 12 and Eq. 13, respectively:

$$W_1 = \left[0.239 \quad 0.352 \quad 0.212 \quad 0.197\right] \tag{12}$$

$$W_2 = \left[0.203 \quad 0.362 \quad 0.217 \quad 0.218\right] \tag{13}$$

The fuzzy membership matrix F of all participants was obtained using the Eq. (9). For example, with the eye response data of 1st participant in high workload level (SI-CHK task), pupil diameter, blink rate, fixation rate and saccade rate were 2.92 mm, 0.39 count/s, 0.82 count/s and 0.47 count/s, respectively. So, the $F_1 st$ shows in Eq.(14):

$$F_1 st = \begin{bmatrix} 0.142 & 0.858 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0.360 & 0.640 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.601 & 0.399 \end{bmatrix} \tag{14}$$

E-MWL score of 1st participant was calculated by Eq.14, 15:

$$B_{1st} = W_1 \times F_{1st} = [0.1068 \quad 0.3131 \quad 0 \quad 0.1307 \quad 0.4494] \tag{15}$$

$$C_{1st} = B_{1st} \times S^T = 70.05 \tag{16}$$

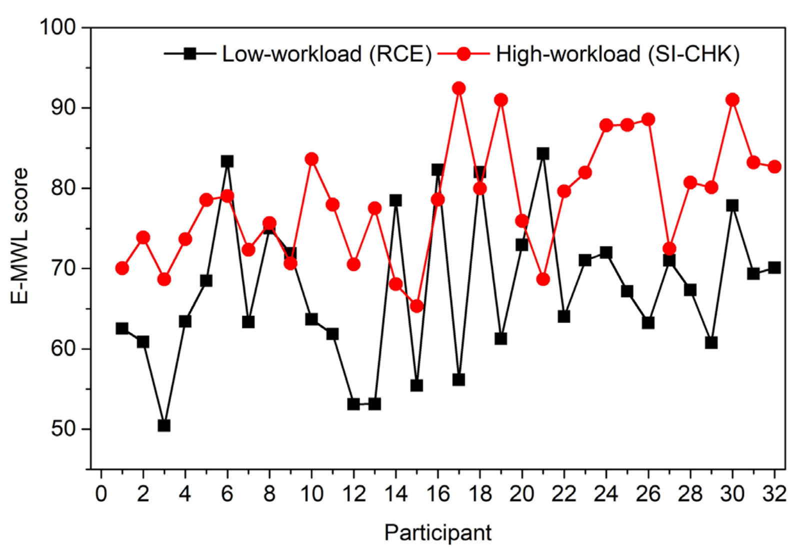
Similarly, the E-MWL score of all participants of two workload levels is plotted in Figure 3.

Validation of E-MWL Method

- Comparison of E-MWL with NASA-TLX.

The correlation analysis (Pearson's correlation) was conducted to investigate the degree of correlation between two variables. The purpose of this step was to validate whether the suggested

Figure 3. The E-MWL score of all participants in low workload (RCE task) and high workload level (SI-CHK task)



E-MWL can measure an operators' MWL. Figure 4 shows the results of a correlation analysis between NASA-TLX and E-MWL scores. Pearson's correlation coefficient between NASA-TLX score and E-MWL is 0.685 (Figure 4a) and 0.752 (Figure 4b) for low and high workload levels, respectively. The results indicated that the correlation between E-MWL and NASA-TLX in high workload is stronger than that between the E-MWL and NASA-TLX in low workload level.

- Comparison of NASA-TLX score and E-MWL with the performance measure indices.

The correlation between NASA-TLX score and the performance measures and that between E-MWL and the performance measures in two workload levels are presented in Figure 5, 6. In the low workload level (Figure 5), the NASA-TLX has a slightly stronger relationship with the operating time than that of the E-MWL; however, the degree of the gap is small. In contrast, the correlation

Figure 4. Correlation between E-MWL and NASA-TLX score of low (a) and high (b) workload level

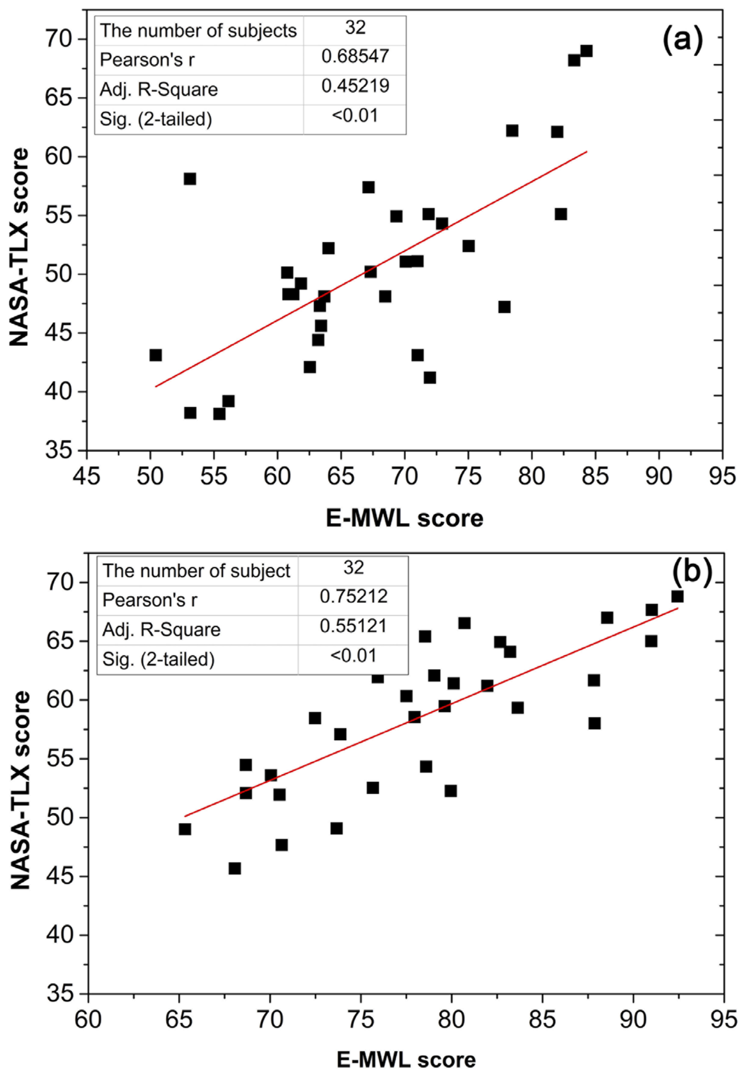
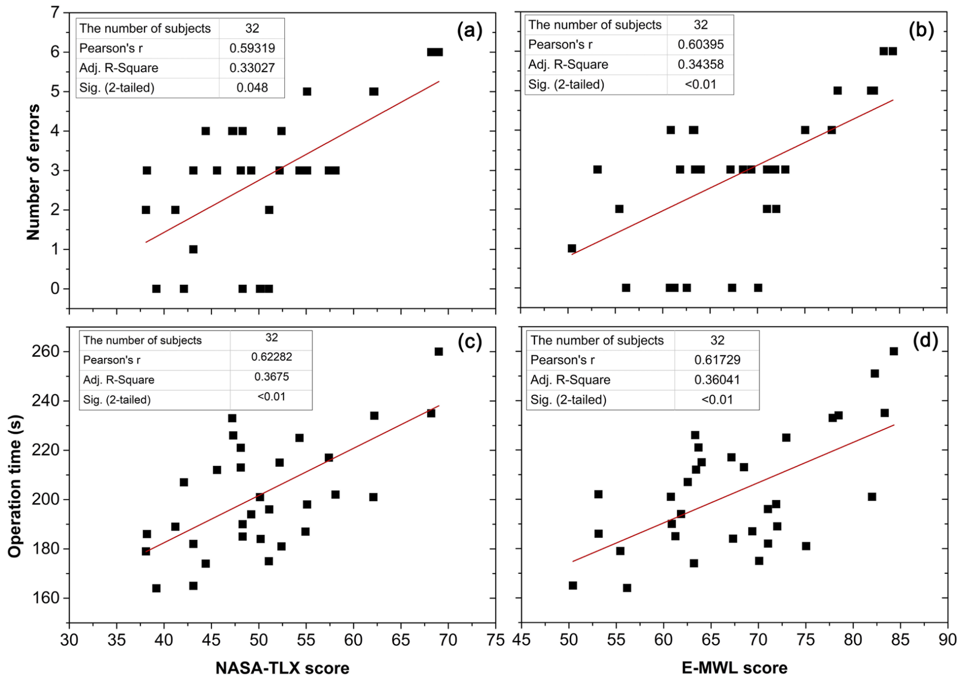


Figure 5. Correlation between NASA-TLX score and performance measures (a, c), E-MWL and performance measures (b, d) in low workload level (RCE task)

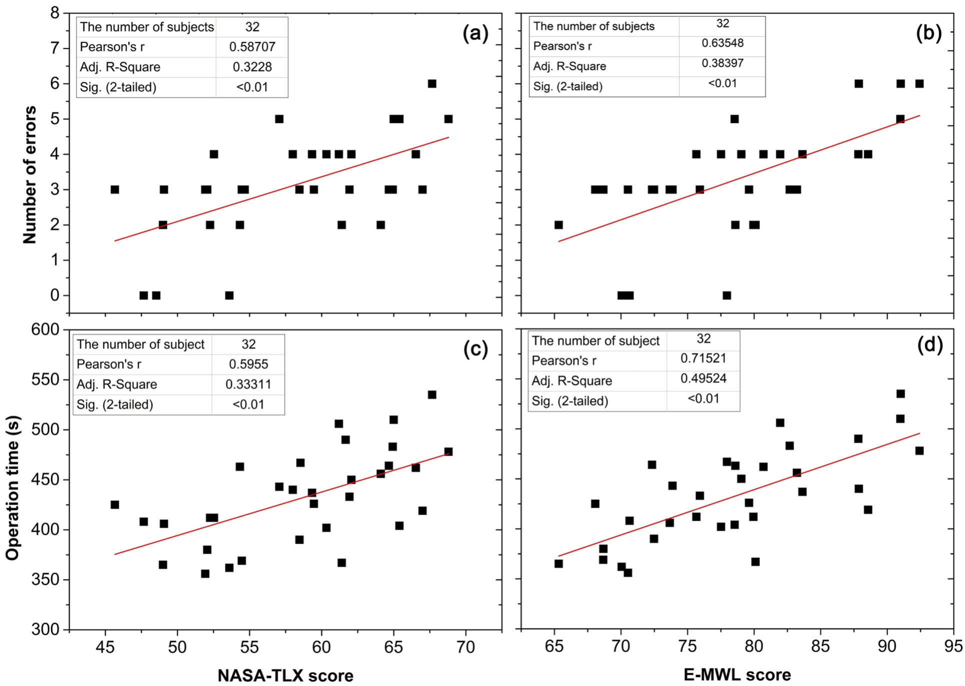


between E-MWL and the number of errors is stronger than that between the NASA-TLX. The analysis data of high workload level shows that E-MWL has a higher correlation with performance measures data than the NASA-TLX score (Figure 6). In addition, the correlation between NASA-TLX score and operation time have an abnormal data deviate from the trend line is greater than the correlation between E-MWL score and the operation time.

DISCUSSION

In this study, the operators' MWL was calculated by fuzzy theory based on the combination of eye response indices. The NASA-TLX scores showed a significant correlation with the different levels of MWL. The operation time and the number of errors also increased as the task complexity increased from easy to difficult. However, no significant difference was found in number of errors. It may be possible to explain this result from the perspective of short operation tasks. A number of previous studies have also concluded that the number of errors is often not sensitive to workload in a short experimental time. The eye response results indicated that pupil diameter increases with the task complexity. Also, there was also a significant difference between the two tasks in blink rate and saccadic rate. This can be attributed to the need for increase processing due to higher workload in the SI-CHK task. The SI-CHK procedure has many operating steps on many different interfaces so operators need to broaden their search to locate and set the required parameters. This result indicates that two tasks used in this experiment could distinguish the different levels of MWL. In addition, the correlation analysis results between the NASA-TLX and eye response indices showed that four eye response indices used in this study were correlated significantly with the NASA-TLX, indicating that these indices may assess and develop the objective MWL evaluation method.

Figure 6. Correlation between NASA-TLX score and performance measures (a, c), E-MWL and performance measures (b, d) in high workload level (SI-CHK task)



E-MWL was developed to measure operators' MWL, and it was validated through experiments by comparison with NASA-TLX and work performance, which has been widely used in NPPs field. The result of the proposed method is the combination with four eye responses indices thanks to the predetermined weights and decreased fuzziness by establishing membership functions. Therefore, this value can ensure reliability without the usual statistical analysis. The correlation analysis results between E-MWL score and NASA-TLX score indicated that Pearson's correlation coefficient is highly correlated, thus, E-MWL can be considered as a tool for measuring the quantity of MWL. The positive correlation between E-MWL data and the NASA-TLX score, indicating that this method reflects the individual's subjective sense of MWL in experimental tasks. In addition, the correlation between the E-MWL and NASA-TLX in high workload level is stronger than that in low workload level, it means that this method is suitable for applying MWL evaluation in high task complexity, and that changes in the complexity of task have a significant influence on subjective ratings.

The validity of the proposed method has been also tested by the correlation between E-MWL and performance measures and that between NASA-TLX and the performance measures in two different tasks. The analysis data of two workload levels show that operation time of performance measures has a slightly higher correlation than the error rate. The reason may be due to error rate is often not sensitive to workload in a short task time as well as in the laboratory environment (Hwang et al., 2008; Lanzetta et al., 1987; Yan et al., 2017). In addition, the data that unusually deviate from the trend line of NASA-TLX more than the E-MWL method. This is due to the subjectivity with which the participants assessed the workload themselves by NASA-TLX method. Even if one had felt uncomfortable, they might have rated it as a low MWL (Choi et al., 2018). Moreover, most subjective rating measures imply (if they do not explicitly state) it is MWL which is being measured and the effects of physical work associated are not considered.

Based on the analysis results, the data of eye response indices can be combined to produce reliable objective MWL in interface control, and the proposed E-MWL is suitable for assessing the objective MWL. Moreover, by comparing the correlation of E-MWL with NASA-TLX, performance measures and eye response indices, E-MWL was shown to be an appropriate objective index for evaluating and predicting operator's MWL. It is of great importance for decision-makers to correctly understand the levels of MWL, and it will be difficult for people who are not acquainted with professional knowledge. Thus, evaluation of operators' MWL objectively with directly measured physiological signals will help the managers to organize the human resources for each task to sustain the appropriate MWL as well as to improve the work performance. However, although physiological signal measurements have the capability to provide us with invaluable information that is just not possible with traditional methods, a major disadvantage is that different people may have different habits and emotional experiences. Thus, it is pertinent that other explicit measures, such as self-report questionnaires or think aloud, be used in conjunction with the physiological response measures.

CONCLUSION

This study develops an objective MWL evaluation method based on eye response data when operating the NPP system via a user interface control, and it was validated through experiments by comparison with the NASA-TLX rating and performance measures method. The proposed E-MWL method was demonstrated to be suitable for evaluating human MWL due to the high correlation with NASA-TLX. By comparing the correlation of E-MWL with the performance measures and that of NASA-TLX with the performance measures data, E-MWL was shown to be suitable for evaluating MWL. This method can be applied to develop an objective evaluation method in the actual work environment to evaluate and compare the effectiveness of different interface designs in NPPs. From this, the developers can allocate human resources for each task to sustain the consistent MWL as well as to improve the interface designs in the system.

However, the proposed E-MWL method has some limitations need to be further investigated. Firstly, this method was validated by students after being trained and practiced. Although they had the knowledge and skills for completing the tasks in the experiment, experience and their psychology might be very different from the experts. Thus, it is necessary to have validation experiments with professional operators or experts. Secondly, E-MWL method was validated through experiment by two tasks with different workload level. However, it is necessary to validate with other experiments with different tasks to confirm the reliability of the proposed method.

ACKNOWLEDGMENT

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

FUNDING AGENCY

The publisher has waived the Open Access Processing fee for this article.

REFERENCES

Ahlstrom, U., & Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics*, *36*(7), 623–636. doi:10.1016/j.ergon.2006.04.002

Asadzadeh, S., Azadeh, A., Negahban, A., & Sotoudeh, A. (2013). Assessment and improvement of integrated HSE and macro-ergonomics factors by fuzzy cognitive maps: The case of a large gas refinery. *Journal of Loss Prevention in the Process Industries*, *26*(6), 1015–1026. doi:10.1016/j.jlp.2013.03.007

Balas, C. E., Koç, M. L., & Tür, R. (2010). Artificial neural networks based on principal component analysis, fuzzy systems and fuzzy neural networks for preliminary design of rubble mound breakwaters. *Applied Ocean Research*, *32*(4), 425–433. doi:10.1016/j.apor.2010.09.005

Batmaz, I., & Ozturk, M. (2008). Using pupil diameter changes for measuring mental workload under mental processing. *Journal of Applied Sciences*, *8*(1), 68–76. doi:10.3923/jas.2008.68.76

Bentivoglio, A. R., Bressman, S. B., Cassetta, E., Carretta, D., Tonali, P., & Albanese, A. (1997). Analysis of blink rate patterns in normal subjects. *Movement Disorder, 12*(6), 1028-1034.

Brookings, J. B., Wilson, G. F., & Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology*, *42*(3), 361–377. doi:10.1016/0301-0511(95)05167-8 PMID:8652753

Casali, J. G., & Wierwille, W. W. (1983). A comparison of rating scale, secondary-task, physiological, and primary-task workload estimation techniques in a simulated flight task emphasizing communications load. *Human Factors*, *25*(6), 623–641. doi:10.1177/001872088302500602 PMID:6671647

Castor, M., Hanson, E., Svensson, E., Nählinder, S., LeBlaye, P., MacLeod, I., Wright, N., Alfredson, J., Ågren, L., & Berggren, P. (2003). *Handbook of mental workload measurement*. GARTEUR.

Čegovnik, T., Stojmenova, K., Jakus, G., & Sodnik, J. (2018). An analysis of the suitability of a low-cost eye tracker for assessing the cognitive load of drivers. *Applied Ergonomics*, *68*, 1–11. doi:10.1016/j.apergo.2017.10.011 PMID:29409621

Chen, W., & Hao, X. (2011). An Optimal Combination Determining Weights Method for Power Quality Fuzzy Synthetic Evaluation. In *Computer Science for Environmental Engineering and EcoInformatics* (pp. 417–423). Springer. doi:10.1007/978-3-642-22694-6_59

Chen, Y., Yan, S., & Tran, C. C. (2019). Comprehensive evaluation method for user interface design in nuclear power plant based on mental workload. *Nuclear Engineering and Technology*, *51*(2), 453–462. doi:10.1016/j.net.2018.10.010

Choi, M. K., Lee, S. M., Ha, J. S., & Seong, P. H. (2018). Development of an EEG-based workload measurement method in nuclear power plants. *Annals of Nuclear Energy*, *111*, 595–607. doi:10.1016/j.anucene.2017.08.032

Chu, W., Li, Y., Liu, C., Mou, W., & Tang, L. (2014). A manufacturing resource allocation method with knowledge-based fuzzy comprehensive evaluation for aircraft structural parts. *International Journal of Production Research*, *52*(11), 3239–3258. doi:10.1080/00207543.2013.869369

Chuang, C.-Y., Lin, C. J., Shiang, W.-J., Hsieh, T.-L., & Liou, J.-L. (2016). Development of an objective mental workload assessment tool based on Rasmussen's skill–rule–knowledge framework. *Journal of Nuclear Science and Technology*, *53*(1), 123–128. doi:10.1080/00223131.2015.1027156

De Waard, D. (1996). *The measurement of drivers' mental workload*. Groningen University, Traffic Research Center Netherlands.

Di Nocera, F., Camilli, M., & Terenzi, M. (2007). A random glance at the flight deck: Pilots' scanning strategies and the real-time assessment of mental workload. *Journal of Cognitive Engineering and Decision Making*, *1*(3), 271–285. doi:10.1518/155534307X255627

Di Stasi, L. L., Antolí, A., & Cañas, J. J. (2013). Evaluating mental workload while interacting with computer-generated artificial environments. *Entertainment Computing*, *4*(1), 63–69. doi:10.1016/j.entcom.2011.03.005

Dyer, R. F., Matthews, J. J., Wright, C. E., & Yudowitch, K. L. (1976). *Questionnaire construction manual*. Army Research Institute.

Gao, Q., Wang, Y., Song, F., Li, Z., & Dong, X. (2013). Mental workload measurement for emergency operating procedures in digital nuclear power plants. *Ergonomics*, *56*(7), 1070–1085. doi:10.1080/00140139.2013.790483 PMID:23654299

Guillon, M., Dumbleton, K., Theodoratos, P., Gobbe, M., Wooley, C. B., & Moody, K. (2016). The effects of age, refractive status, and luminance on pupil size. *Optometry and Vision Science*, *93*(9), 1093–1100. doi:10.1097/OPX.0000000000000893 PMID:27232893

Hampson, R., Opris, I., & Deadwyler, S. (2010). Neural correlates of fast pupil dilation in nonhuman primates: Relation to behavioral performance and cognitive workload. *Behavioural Brain Research*, *212*(1), 1–11. doi:10.1016/j.bbr.2010.03.011 PMID:20226215

Hsieh, M.-C., Chiu, M.-C., & Hwang, S.-L. (2015). An optimal range of information quantity on computer-based procedure interface design in the advanced main control room. *Journal of Nuclear Science and Technology*, *52*(5), 687–694. doi:10.1080/00223131.2014.977832

Hwang, S.-L., Yau, Y.-J., Lin, Y.-T., Chen, J.-H., Huang, T.-H., Yenn, T.-C., & Hsu, C.-C. (2008). Predicting work performance in nuclear power plants. *Safety Science*, *46*(7), 1115–1124. doi:10.1016/j.ssci.2007.06.005

Iqbal, S. T., Zheng, X. S., & Bailey, B. P. (2004). Task-evoked pupillary response to mental workload in human-computer interaction. In *CHI '04 Conference on Human Factors in Computing Systems* (pp. 1477-1480). Association for Computing Machinery (ACM). doi:10.1145/985921.986094

Isaac, A., Shorrock, S. T., & Kirwan, B. (2002). Human error in European air traffic management: The HERA project. *Reliability Engineering & System Safety*, *75*(2), 257–272. doi:10.1016/S0951-8320(01)00099-0

Jiang, F., Zheng, Q., & Shi, W. (2012). The applied research of fuzzy comprehensive evaluation on talent training mode of safety engineering. *Procedia Engineering*, *43*, 425–430. doi:10.1016/j.proeng.2012.08.073

Jiao, J., Ren, H., & Sun, S. (2016). Assessment of surface ship environment adaptability in seaways: A fuzzy comprehensive evaluation method. *International Journal of Naval Architecture and Ocean Engineering*, *8*(4), 344–359. doi:10.1016/j.ijnaoe.2016.05.002

Jou, Y.-T., Yenn, T.-C., Lin, C. J., Yang, C.-W., & Chiang, C.-C. (2009). Evaluation of operators' mental workload of human-system interface automation in the advanced nuclear power plants. *Nuclear Engineering and Design*, *239*(11), 2537–2542. doi:10.1016/j.nucengdes.2009.06.023

Kang, J., Zhang, J., & Gao, J. (2016). Improving performance evaluation of health, safety and environment management system by combining fuzzy cognitive maps and relative degree analysis. *Safety Science*, *87*, 92–100. doi:10.1016/j.ssci.2016.03.023

Kankal, M., & Yüksek, Ö. (2012). Artificial neural network approach for assessing harbor tranquility: The case of Trabzon Yacht Harbor, Turkey. *Applied Ocean Research*, *38*, 23–31. doi:10.1016/j.apor.2012.05.009

Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. *Multiple-Task Perform*, (7), 279–328.

Kum, S., Furusho, M., Duru, O., & Satir, T. (2007). Mental workload of the VTS operators by utilising heart rate. *International Journal on Marine Navigation and Safety of Sea Transportation*, *1*(2), 145–151.

Lanzetta, T. M., Dember, W. N., Warm, J. S., & Berch, D. B. (1987). Effects of task type and stimulus heterogeneity on the event rate function in sustained attention. *Human Factors*, *29*(6), 625–633. doi:10.1177/001872088702900602 PMID:3436619

Lehrer, P., Karavidas, M., Lu, S. E., Vaschillo, E., Vaschillo, B., & Cheng, A. (2010). Cardiac data increase association between self-report and both expert ratings of task load and task performance in flight simulator tasks: An exploratory study. *International Journal of Psychophysiology*, *76*(2), 80–87. doi:10.1016/j.ijpsycho.2010.02.006 PMID:20172000

Li, D., Sui-huai, Y., & Wen-jun, W. (2013). Research on Fuzzy Comprehensive Evaluation of Human-Machine Interface Layout of Driller Control Room. In Computational and Information Sciences (ICCIS) (pp. 1114-1117). IEEE. doi:10.1109/ICCIS.2013.296

Nachreiner, F. (1995). Standards for ergonomics principles relating to the design of work systems and to mental workload. *Applied Ergonomics*, *26*(4), 259–263. doi:10.1016/0003-6870(95)00029-C PMID:15677027

Naderpour, M., Lu, J., & Zhang, G. (2016). A safety-critical decision support system evaluation using situation awareness and workload measures. *Reliability Engineering & System Safety*, *150*, 147–159. doi:10.1016/j.ress.2016.01.024

Nakayama, M., Takahashi, K., & Shimizu, Y. (2002). 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* (pp. 37-42). Association for Computing Machinery (ACM). doi:10.1145/507072.507080

Nyström, M., Hooge, I., & Andersson, R. (2016). Pupil size influences the eye-tracker signal during saccades. *Vision Research*, *121*, 95–103. doi:10.1016/j.visres.2016.01.009 PMID:26940030

Orlandi, L., & Brooks, B. (2018). Measuring mental workload and physiological reactions in marine pilots: Building bridges towards redlines of performance. *Applied Ergonomics*, *69*, 74–92. doi:10.1016/j.apergo.2018.01.005 PMID:29477333

Pan, B., Hembrooke, H. A., Gay, G. K., Granka, L. A., Feusner, M. K., & Newman, J. K. (2004). The determinants of web page viewing behavior: an eye-tracking study. In *Proceedings of the 2004 symposium on Eye tracking research & applications* (pp. 147-154). Association for Computing Machinery (ACM) doi:10.1145/968363.968391

Peshkovskaya, A. G., Babkina, T. S., Myagkov, M. G., Kulikov, I. A., Ekshova, K. V., & Harriff, K. (2017). The socialization effect on decision making in the Prisoner's Dilemma game: An eye-tracking study. *PLoS One*, *12*(4), 1–15. doi:10.1371/journal.pone.0175492 PMID:28394939

Philippe, R., Christine, C., Chiara, N., Samantha, J., & Toumelin, N. (2016). Analysis of submarine steering: Effects of cognitive and perceptual–motor requirements on the mental workload and performance of helmsmen. *Cognition Technology and Work*, *18*(4), 657–672. doi:10.1007/s10111-016-0384-0

Proctor, R. W., & Van Zandt, T. (2018). *Human factors in simple and complex systems*. CRC Press.

Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimental Psychology. Applied*, *6*(1), 31–43. doi:10.1037/1076-898X.6.1.31 PMID:10937310

Rosch, J. L., & Vogel-Walcutt, J. J. (2013). A review of eye-tracking applications as tools for training. *Cognition Technology and Work*, *15*(3), 313–327. doi:10.1007/s10111-012-0234-7

Ryu, K., & Myung, R. (2005). Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *International Journal of Industrial Ergonomics*, *35*(11), 991–1009. doi:10.1016/j.ergon.2005.04.005

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x

Sheridan, T., & Stassen, H. (1979). Definitions, models and measures of human workload. In *Mental Workload*. Springer. doi:10.1007/978-1-4757-0884-4_12

Sibert, L. E., & Jacob, R. J. (2000). Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 281-288). Association for Computing Machinery (ACM). doi:10.1145/332040.332445

Takuska-Węgrzyn, E. (2008). Application of statistical methods for evaluation of rock-burst risks in copper ore mine conditions. *Archives of Mining Sciences*, *53*(1), 23–30.

Tran, C. C., Yan, S., Habiyaremye, J. L., & Wei, Y. (2017). Predicting driver's work performance in driving simulator based on physiological indices. In *International Conference on Intelligent Human Computer Interaction* (pp. 150-162). Springer. doi:10.1007/978-3-319-72038-8_12

Tsang, P. S., & Vidulich, M. A. (2006). Mental workload and situation awareness (3rd ed). Wiley.

Van Orden, K. F., Jung, T.-P., & Makeig, S. (2000). Combined eye activity measures accurately estimate changes in sustained visual task performance. *Biological Psychology*, *52*(3), 221–240. doi:10.1016/S0301-0511(99)00043-5 PMID:10725565

Veltman, J., & Gaillard, A. (1998). Physiological workload reactions to increasing levels of task difficulty. *Ergonomics*, *41*(5), 656–669. doi:10.1080/001401398186829 PMID:9613226

Wang, T.-C., & Lee, H.-D. (2009). Developing a fuzzy TOPSIS approach based on subjective weights and objective weights. *Expert Systems with Applications*, *36*(5), 8980–8985. doi:10.1016/j.eswa.2008.11.035

Wang, Y., Li, Y., Liu, W., & Gao, Y. (2015). Assessing operational ocean observing equipment (OOOE) based on the fuzzy comprehensive evaluation method. *Ocean Engineering*, *107*, 54–59. doi:10.1016/j.oceaneng.2015.07.032

Wang, Y., Toor, S. S., Gautam, R., & Henson, D. B. (2011). Blink frequency and duration during perimetry and their relationship to test–retest threshold variability. *Investigative Ophthalmology & Visual Science*, *52*(7), 4546–4550. doi:10.1167/iovs.10-6553 PMID:21447676

Wierda, S. M., van Rijn, H., Taatgen, N. A., & Martens, S. (2012). Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(22), 8456–8460. doi:10.1073/pnas.1201858109 PMID:22586101

Wilson, G., Schlegel, R., & Veltman, J. (2004). *Operator functional state assessment.* Report No. RTO-TR-HFM-104 AC/323 (HFM-104) TP/48 RTO. N. R. a. T. O. RTO.

Wilson, G. F. (2002). An analysis of mental workload in pilots during flight using multiple psychophysiological measures. *The International Journal of Aviation Psychology*, *12*(1), 3–18. doi:10.1207/S15327108IJAP1201_2

Wu, X., & Li, Z. (2013). Secondary task method for workload measurement in alarm monitoring and identification tasks. In *International Conference on Cross-Cultural Design* (pp. 346-354). Springer. doi:10.1007/978-3-642-39143-9_39

Wyatt, H. J. (1995). The form of the human pupil. *Vision Research*, *35*(14), 2021–2036. doi:10.1016/0042-6989(94)00268-Q PMID:7660606

Xie, B., & Salvendy, G. (2000). Review and reappraisal of modelling and predicting mental workload in single- and multi-task environments. *Work and Stress*, *14*(1), 74–99. doi:10.1080/026783700417249

Xie, Q., Ni, J.-Q., & Su, Z. (2017). Fuzzy comprehensive evaluation of multiple environmental factors for swine building assessment and control. *Journal of Hazardous Materials*, *340*, 463–471. doi:10.1016/j.jhazmat.2017.07.024 PMID:28759867

Yan, S., Tran, C. C., Chen, Y., Tan, K., & Habiyaremye, J. L. (2017). Effect of user interface layout on the operators' mental workload in emergency operating procedures in nuclear power plants. *Nuclear Engineering and Design*, *322*, 266–276. doi:10.1016/j.nucengdes.2017.07.012

Yan, S., Tran, C. C., Wei, Y., & Habiyaremye, J. L. (2019). Driver's mental workload prediction model based on physiological indices. *International Journal of Occupational Safety and Ergonomics*, *25*(3), 476–484. doi:10.1080/10803548.2017.1368951 PMID:28820660

Yang, Q. (2010). State Oriented Procedure Application in Nuclear Power Plants in China. In *International Conference on Nuclear Engineering* (pp. 183-186). ASME. doi:10.1115/ICONE18-29591

Zadeh, L. A. (1965). Information and control. *Fuzzy Sets, 8*(3), 338-353.

Zhou, J., Li, X., & Shi, X. (2012). Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines. *Safety Science*, *50*(4), 629–644. doi:10.1016/j.ssci.2011.08.065

Zou, Z.-H., Yun, Y., & Sun, J.-N. (2006). Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment. *Journal of Environmental Sciences (China)*, *18*(5), 1020–1023. doi:10.1016/S1001-0742(06)60032-6 PMID:17278765

Cong Chi Tran is a lecturer at the College of Electromechanical and Civil Engineering, Vietnam National University of Forestry, Viet Nam. He lectures and consults on a variety of subjects on mechanical engineering technology. His current research interests include electromechanical in Agro-forestry, human-computer interaction, and human factors engineering for safety-critical systems. His research articles have appeared in such journals as Nuclear Engineering and Design, International Journal of Industrial Ergonomics, Nuclear Technology among others. He received his BS.c (2007) and MS.c (2011) degree in Mechanical Engineering at the Hanoi University of Science and Technology, Viet Nam. He held his Ph.D. degree in Mechanical Engineering from Harbin Engineering University, China in 2019.

Shengyuan Yan is a professor at the College of Mechanical and Electrical Engineering, Harbin Engineering University, China. He received the BS.c (1984), MS.c (1988), and Ph.D. (2005) in Harbin Engineering University. He lectures and consults on a variety of subjects on human factors engineering and human-machine interface optimization. He has published many articles in prestigious journals such as Nuclear Engineering and Design, Nuclear Engineering and Technology, Nuclear Technology, IEEE Access, and others. His current research interests include designing a three-dimensional simulation system for Human-Machine Environment, evaluation and optimization design for Ships, nuclear power plants, and other industrial plants.