ERGONOMICS
https://doi.org/10.1080/00140139.2022.2134590

Taylor & Francis
Taylor & Francis Group

Check for updates

ARTICLE

Influence of nuclear power plant interface complexity on user decision-making: an ERP study

Wenzhe Tang, Shanguang Chen, Chengqi Xue, Yafeng Niu, Junkai Shao and Yanfei Zhu

School of Mechanical Engineering, Southeast University, Nanjing, China

ABSTRAC

User decision-making concerning critical operations is very important to nuclear power plant (NPP) safety. The NPP interface is the main information source that guides decision-making; thus, a good interface design is essential. Among the interface design factors such as interface complexity, layout and colour, interface complexity (the amount of information in the interface) has the greatest impact on NPP operator decision-making. This paper used the event-related potential (ERP) to evaluate the impact of interface complexity on user decision-making and found interface complexity has a specific range suitable for decision-making. Based on this important finding, a fast and economical method of evaluating NPP interfaces in all design phases was proposed. This method compensates for the shortcomings of traditional methods, such as heuristic evaluation and experimental evaluation, which are inconvenient for evaluating interfaces in initial design phase; it can also be applied to interfaces with similar features in other industrial fields.

Practitioner summary: Evaluation of the impact of NPP interface complexity on user decision-making through an ERP experiment revealed a specific range of interface complexity that facilitates user decision-making. Based on this finding, a new, fast and inexpensive interface evaluation method was proposed.

Abbreviations: NPP: nuclear power plant, it is a thermal power station in which the heat source is a nuclear reactor; ERP: event-related potential, it is the measured brain response that is the direct result of a specific cognitive, or motor event.

ARTICLE HISTORY
Received 25 September 2020
Accepted 3 October 2022

KEYWORDS
Nuclear power plant; interface complexity; decision-making; ERP; evaluation method

1. Introduction

In modern nuclear power plants (NPPs), control systems generally operate automatically; thus, the operator's job has changed from controlling to monitoring and decision-making (Hugo and Gertman 2016). Human decision-making is currently the main factor that affects the safety of NPPs (Bohanec et al. 2020). Decision-making is the process of evaluating and selecting an option from different alternatives (Shukla, Auriol, and Hipel 2016). It is affected by many factors, such as information, environment, and personality. For NPP operators, decision-making is the process of selecting different feedback operations based on the information obtained from the interface. Since the environment of the main control room is relatively stable and the operators have been trained, interface information is the main uncontrollable decision-making factor. The interface plays an important role in NPP safety, integrating almost all alarms and other information. Poor interface information will lead to incorrect judgements by the operator, resulting in serious consequences.

Existing research on the relationship between interfaces and decision-making focuses on interface design elements such as layout, colour and font (Starke and Baber 2018; Oyibo and Vassileva 2020). The layout is closely related to functional division and information distribution and therefore has an impact on user decision-making. Among the various interface layouts, such as vertical, horizontal, centre, and off-centre, the vertical layout is more conducive to user decision-making than the horizontal layout (Chen, Li, and Jamieson 2018), and the centre layout is better than the off-centre layout (Chen et al. 2021). Colour can attract the user's attention and affect the user's mood. For example, research has found that the background colour of a web page has an impact on user decision-making. Cool colours are more conducive to user decision-making than warm colours (Cheng, Wu, and Leiner 2019), and high saturation (blue) is more

CONTACT Chengqi Xue ipd_xcq@seu.edu.cn School of Mechanical Engineering, Southeast University, Nanjing, China

conducive to user decision-making than low saturation (grey) (Westerman et al. 2012). The legibility of font also influences the acquisition of information and is closely related to user decision-making bias. Among the various fonts used to present information on the interface, prominent fonts and fonts that are difficult to read can reduce user decision-making bias, thereby facilitating decision-making (Shen and Urminsky 2013; Korn et al. 2018; Diaz-Lago and Matute 2019).

The literature contains a large number of papers on the relationship between digital interfaces and decision-making. However, the types of digital interfaces are limited, mainly to web pages and personal computer interfaces. NPP interfaces are special interfaces that present a large amount of information as well as multiple sources of information. Such interfaces are also widely used in complex information systems such as aeroplanes and ships, which require high safety and efficiency. Unlike web pages and personal computer interfaces, NPP interfaces have limited colour and layout options, and interface design is restricted by safety regulations. In addition, most studies have focussed on the impact of local interface elements such as layout, colour and font on decision-making, but research on global design factors, such as interface complexity, is lacking. Interface complexity refers to the overall amount of information provided by an interface and is considered to have a large impact on decision-making concerning complex information system interfaces (Maglic and Zec 2020). However, findings on the relationship between interface complexity and decision-making are inconsistent. Some studies have reported that interface complexity is negatively related to user decision-making; that is, the lower the interface complexity is, the better the user decision-making (Petrovcic et al. 2018; Vincent et al. 2019; Guo et al. 2021). Others have found that interface complexity is positively related to user decision-making; that is, the more information the interface provides, the better the user decision-making (Han, Xue, and Zhang 2017; Lazard and King 2020). Regardless, interface complexity that is too high or too low seems to be detrimental to user decision-making. Further experimental evidence is needed to determine whether a specific range of interface complexity facilitates decision-making.

The aim of this paper was to evaluate the influence of NPP interface complexity on user decision-making and to experimentally determine the upper and lower limits of interface complexity. Three interfaces with different interface complexity were designed based on the process display interfaces of NPPs. User behavioural and event-related potential (ERP) data were recorded and analysed. The study focussed on determining whether an interface complexity threshold affects decision-making. The insights provided will deepen current understanding of the impact of interface complexity on decision-making. Additionally, a practical motivation for this study was to propose an evaluation method suitable for the initial design phase of an interface, which is highly important for interface designers.

2. Methods

NPP operator decision-making involves rapid decisions based on the operator's interaction with the interface; thus, a method with high temporal resolution was needed. Electroencephalography (EEG) has natural advantages. It has extremely high temporal resolution, can assess responses triggered by events (event-related potentials; ERP), and directly reflects changes in the brain (Li et al. 2018; Zhang 2018; Changoluisa, Varona, and De Borja Rodriguez 2020). Ba et al. (2016) used ERP to study driver decision-making, and Meng and Xiu (2018) used ERP to study risk-related decision-making. The use of ERP is thus an effective research method and was adopted in this study.

The overall design of the experiment is discussed here. The experiment adopted a within-participants design and consisted of two parts: familiarisation and test phases. The participants were instructed to select different feedback operations based on the information provided by the interface. The detailed schedules of each part are provided in detail later. The experimental schedule for each participant is shown in Figure 1.

Decision-making tasks mainly elicit activity in the frontal, temporal, and parietal lobes of the brain

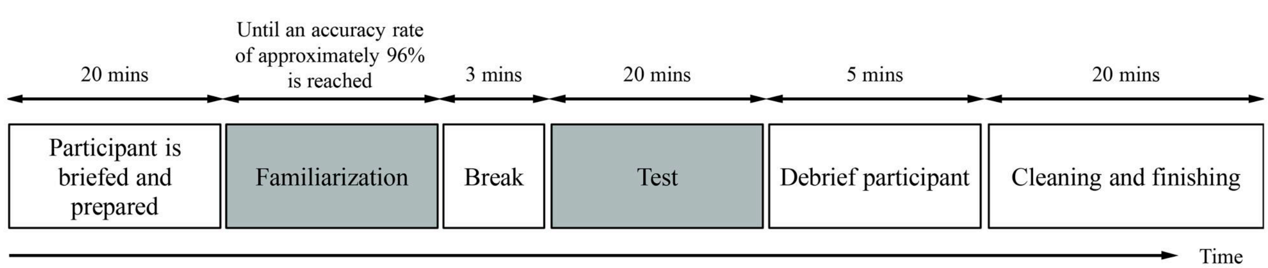


Figure 1. Experimental schedule for each participant.

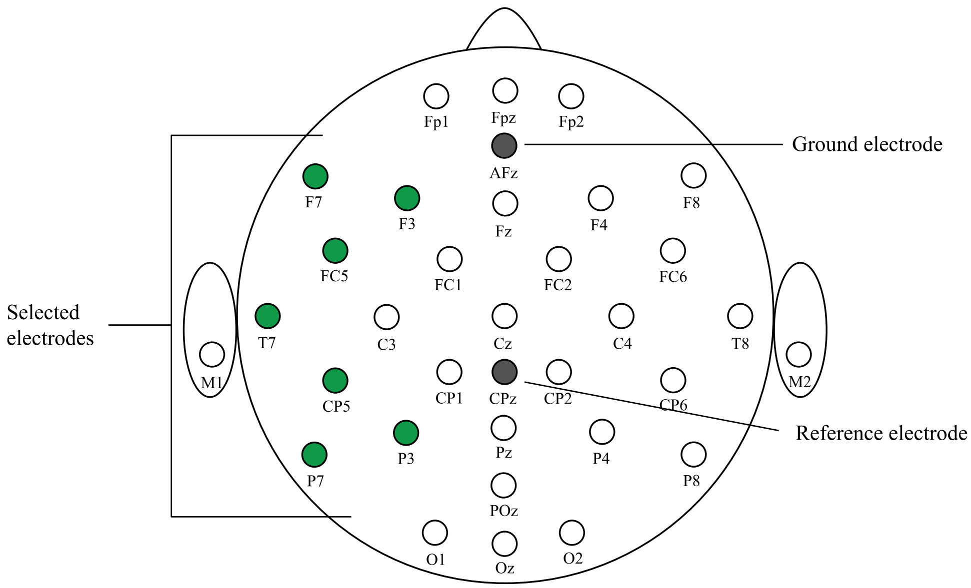


Figure 2. Electrode placement layout.

(Ahmad et al. 2016; Sandor et al. 2018; Guidotti et al. 2019). Overall, activation is stronger in the left hemisphere, though this depends on the type of decision-making task employed (Ernst et al. 2004). Therefore, in the present study, 7 electrodes in the left frontal lobe, left temporal lobe and left parietal lobe at sites F7, F3, FC5, T7, P7, CP5 and P3 were selected, as shown in Figure 2. The EEG waveform changes within 200 ms before initiation of the key press, the average visual reaction time window (Collins, Abbott, and Richards 2011); in this manner, the waveform elicited during the decision-making period could be examined. The behavioural data included the accuracy rate (ACC) and reaction time (RT), and the ERP data included the amplitude and latency.

2.1. Participants

Since the experiment aimed to examine differences in human decision-making depending on interface complexity, prior experience with operating NPP interfaces was not required. The lack of previous experience operating NPP interfaces (i.e. naive participants) was addressed during the familiarisation phase in which each participant was allotted sufficient time to become fully accustomed to the experimental NPP interfaces (see '2.4 Procedure'). A total of 56 participants were enrolled in this experiment and their characteristics are shown in Table 1.

Table 1. Participant characteristics.

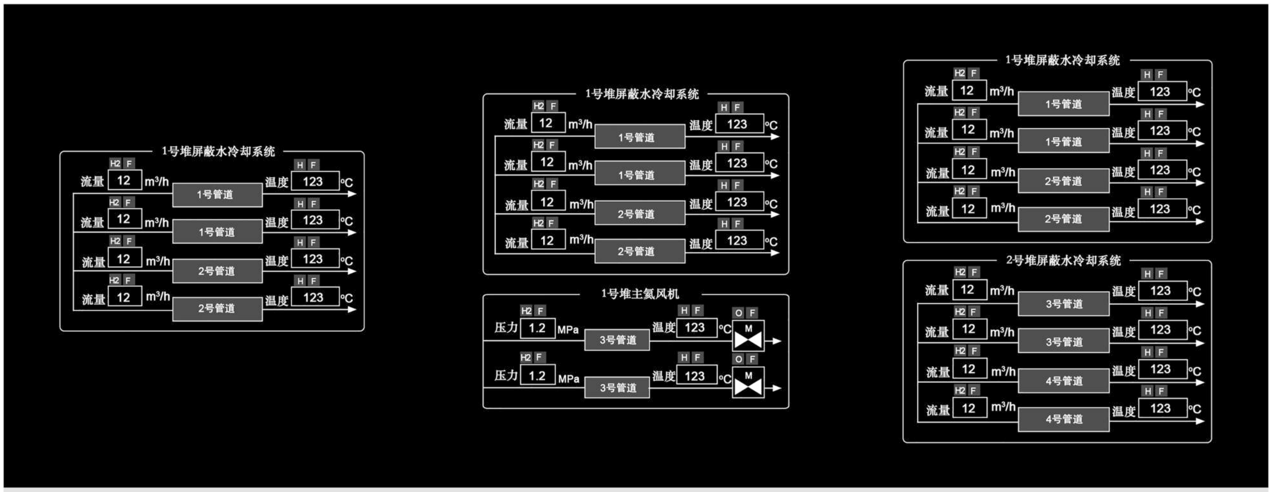
Profile	56 Graduate students
Sex	34 Males, 22 females
Age	Ranging from 21 to 30 years old, with an average age of 25 years
Handedness	All right-handed

2.2. Stimuli

The experimental stimuli were designed based on the process display interfaces of NPPs. Interface complexity can be measured in terms of image entropy, which represents the average amount of information provided in an image. Image entropy can be expressed by the following formula:

$$H = -\sum_{i=1}^{L} p_i \log_2 p_i$$

where H is the image entropy, p_i is the probability of a certain grey level in the entire image, and L is the total number of grey levels in the image. Interface complexity can be measured by image entropy for three reasons. First, the NPP interface was predesigned and could not be created or closed by the user. Second, the number of colours was limited, with generally no more than five colours, and the contrast between colours was large enough for an image to retain the grey level and its information after greyscale processing. Third, the image was converted from a graphic; thus, there was no noise, and each pixel was independent. The general procedure for calculating the image entropy of the interface is as follows: first, the interface was converted into an 8-bit grey image (reducing the calculation cost compared to that for colourful images) and then the probability distribution of the grey levels in the image and the image entropy were calculated according to the formula.



(a) Low complexity
Image entropy: 0.6971

(b) Moderate complexity
Image entropy: 1.0324

(c) High complexity
Image entropy: 1.3637

Figure 3. Experimental stimuli with varying levels of interface complexity.

Table 2. Eego mylab system parameters.

Amplifier size	160 × 205 × 22 mm	Amplifier weight	500 g
Cap electrodes	32	Electrode material	Ag/AgCl
Input resistance	1 GΩ	Input signal range	150 ~ 1000 mVPP
Sampling frequency	16 kHz	Interface	USB, TTL

All stimuli were decolourised, and an intermediate layout was adopted to prevent colour, layout and other factors from interfering with the ERP waveform. This approach was appropriate because the colour information was converted into grey level information, and the layout had little effect on the amount of interface information. The interface complexity of the stimuli increased proportionally, with three levels of image entropy, labelled low, moderate, and high complexity, as shown in Figure 3. The resolution of each stimulus was 1024 × 768 pixels.

2.3. Apparatus

The experiment was conducted in the Human Factors Laboratory. The ERP equipment was the eego mylab system (ANT Neuro, the Netherlands), and the system parameters are shown in Table 2. The eego mylab system included a laptop with eego recording and analysis software for ERP signal recording and analysis. A desktop computer with E-Prime 2.0 software was used for task presentation. The screen size of the desktop computer was 23.8 inches, and the resolution was 1024 × 768 pixels.

2.4. Procedure

This research was approved by the clinical research Independent Ethics Committee of Zhongda Hospital affiliated with Southeast University (2021ZDSYLL201-P01). Participants received a briefing document approximately three days before the experiment that explained the purpose of the experiment, the tasks to be performed and the scheduled timetable. Each participant provided informed consent. In the first step of the experiment, after participants entered the laboratory on the day of the experiment, they cleaned and dried their hair. Then, the experimenter helped the participant put on the EEG cap and used EEG gel to connect cap electrodes to the participant's scalp. The impedance of each electrode was kept below 5 kΩ. The ground electrode was placed at AFz, and the reference electrode was placed at CPz, as shown in Figure 2. A sampling frequency of 1000 Hz was used in the experiment. Next, participants placed their hands on the keyboard in a natural position and maintained a distance of 55-65 cm between their eyes and the monitor screen. In the experiment, participants were instructed to avoid unnecessary body movements, such as moving their

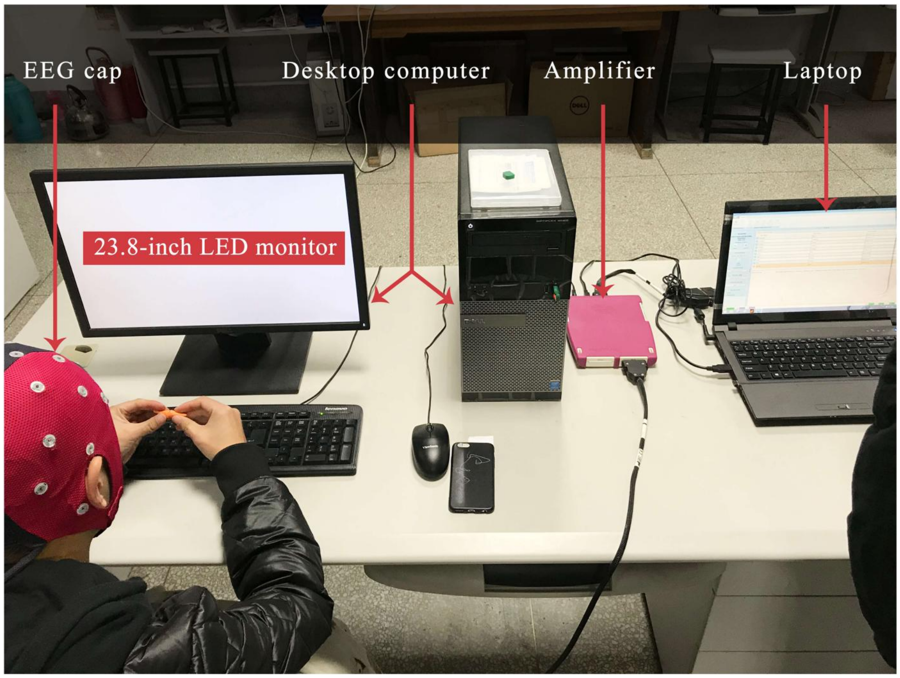


Figure 4. The eego mylab system.

legs or head. The eego mylab system is shown in Figure 4.

2.4.1. Familiarisation phase

In familiarisation phase of the experiment, participants learned the correct key press operations for different situations through feedback. Interfaces without abnormal data required participants to press the space key, while interfaces with abnormal data required participants to press number keys representing areas where abnormal data were shown. Most importantly, in this phase of the experiment, participants were asked to emphasise both speed and accuracy to achieve the speed-accuracy trade-off that is necessary for NPP operator decision-making. Real-time feedback was provided on the ACC and RT of the participant during all practice runs. When participants reached an ACC of approximately 96% (Van Zon et al. 2020), they were considered to be well trained. A flow chart of the familiarisation phase is shown in Figure 5(a).

2.4.2. Test phase

In test phase, participants followed the instructions on the screen. A total of 120 interfaces were displayed in a random order, with each of the three image entropy levels appearing 40 times. The stimuli were divided into 4 blocks, and each block consisted of 30 trials. In each trial, a fixation cross (+) was presented at the centre of the screen for 500 ms, and then an interface appeared. The participant was instructed to observe the interface, make a decision as soon as possible and press one of the corresponding keys, which were the same keys used during the familiarisation phase. Subsequently, the ACC of their decision was then displayed for 500 ms to provide feedback on their speed-accuracy trade-off. E-prime software and eego software simultaneously recorded the behavioural and ERP data, respectively, throughout the experiment. A flow chart of the test phase of the experiment is shown in Figure 5(b).

After the experiment, participants answered questions about their experience and challenges encountered in the experiment. Each participant received a gift in appreciation of their participation.

3. Results

All 56 participants reached the target ACC of 96% and passed the familiarisation phase of the experiment, but in the test phase, 2 participants were excluded due to their obvious drift in ERP data. Thus, valid data was obtained from a total of 54 participants.

3.1. Behavioural data

Behavioural data included the ACC and RT of participants' decision-making outcomes. The ACC was calculated as the ratio of the number of trials with correct keystrokes to the total number of trials. The RT was defined as the average duration between the appearance of the image to the time when a key was

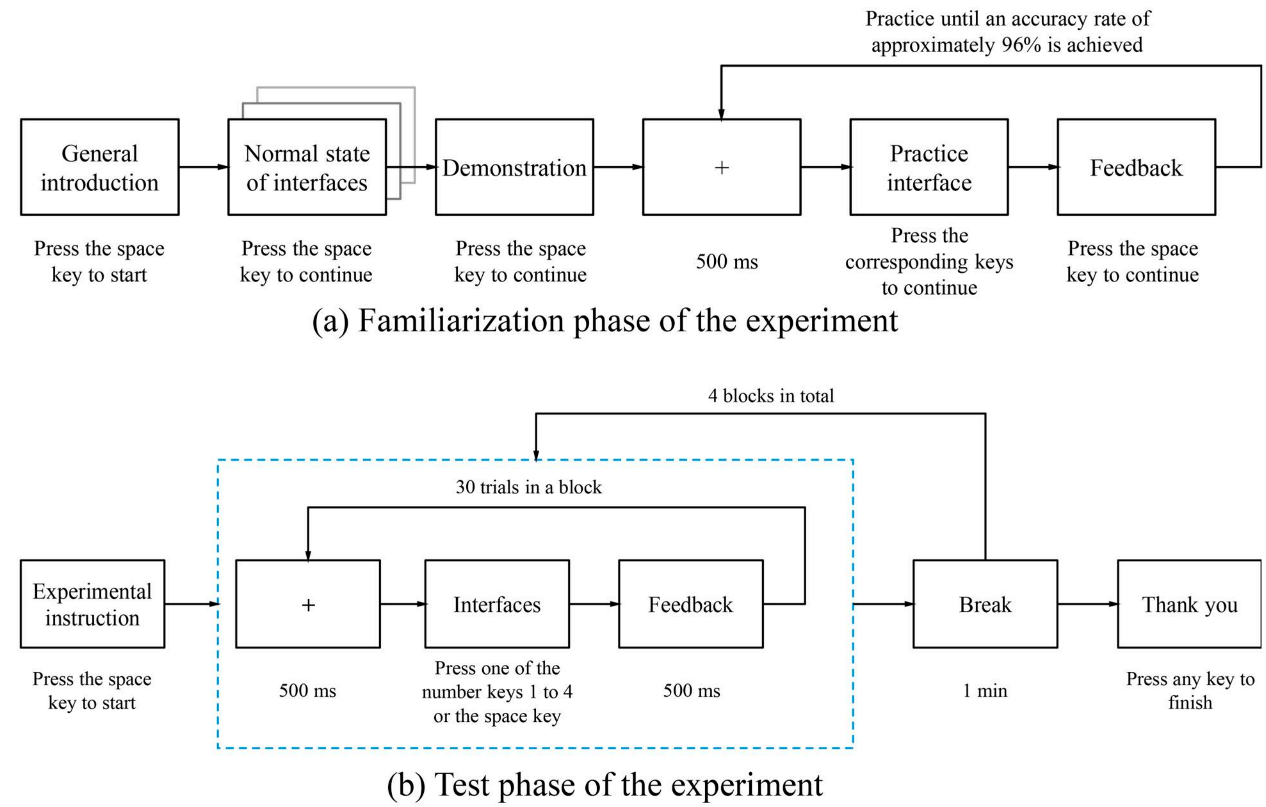


Figure 5. Flow chart of the experimental procedure.

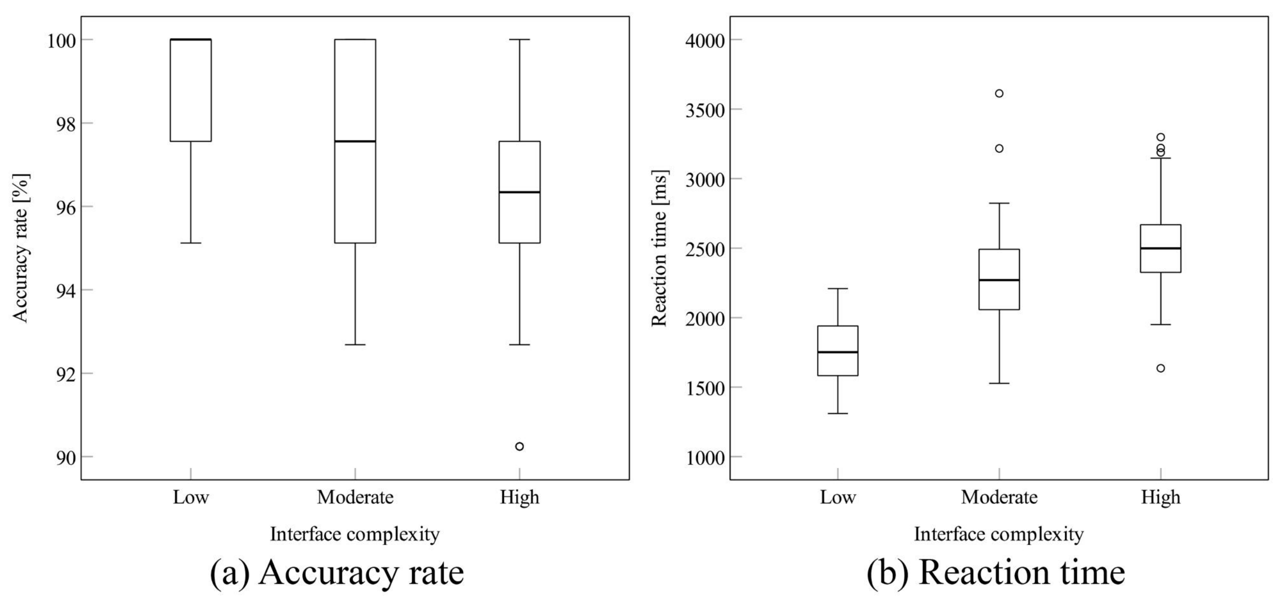


Figure 6. Box plots of decision-making behavioural data.

pressed. Box plots of the decision-making behavioural data are shown in Figure 6. Figure 6(a) shows the ACC according to interface complexity: the median ACC (from highest to lowest) followed the order low, moderate, and high complexity; the mean ACC also followed this order. Paired-sample *t* tests were performed to analyse the results, and the significance level p was set to 0.05. The results showed that the average ACC between low and moderate complexity ($p < 0.05$), between low and high complexity ($p < 0.001$), and between moderate and high complexity ($p < 0.05$) significantly differed. Figure 6(b) shows the RT according to interface complexity: the median RT (from highest to lowest) followed the order high,

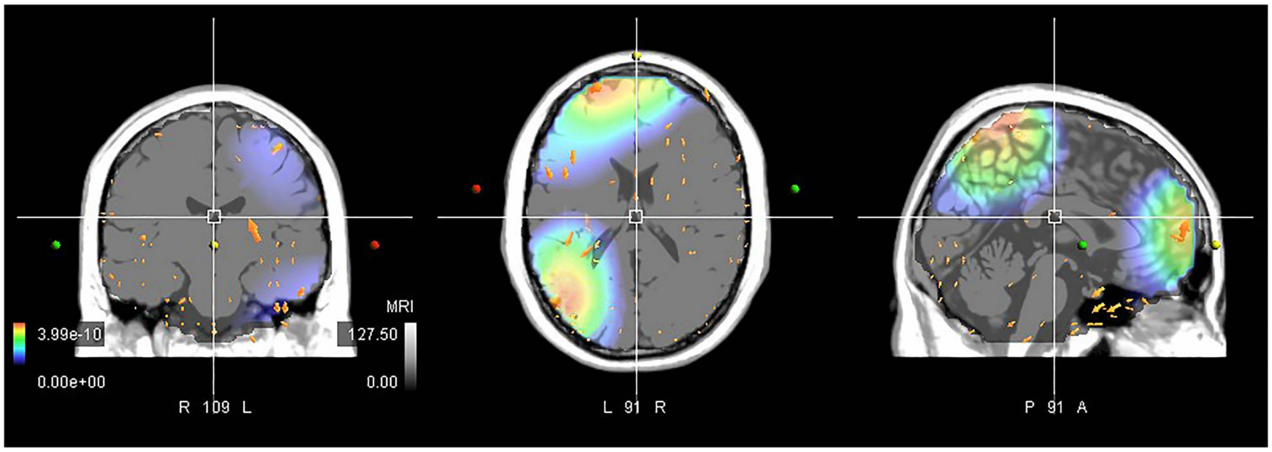


Figure 7. Activated brain regions according to source analysis.

moderate, and low complexity; the mean RT also followed this order. Paired-sample t tests were performed to analyse the results, and the significance level p was set to 0.05. The results showed that the average RT between low and moderate complexity ($p < 0.001$), between low and high complexity ($p < 0.001$), and between moderate and high complexity ($p < 0.001$) significantly differed. Thus, interface complexity has a significant impact on decision-making.

3.2. ERP data

The raw EEG data were processed by ASA software (ANT Neuro, the Netherlands). The EEG signals were bandpass filtered (0.1–40 Hz) and digitised at 1000 Hz. The EEG was re-referenced to mastoids, corrected for blink artefacts using independent component analysis, and low pass filtered at 30 Hz. Baseline corrections were made for EEG epochs between −400 and 200 ms after key press by subtracting the average voltage during the 200 ms period after key press and segments containing residual artefacts exceeding ±70 μV (peak-to-peak) were excluded. The corrected EEG epochs were averaged separately for each participant and interface complexity level, and finally a grand averaging was performed for all participants.

The processed ERP data showed that under different interface complexity conditions, the left frontal lobe, left temporal lobe and left parietal lobe all exhibited obvious negative waves. Source analysis revealed that the activated brain regions were mainly located in the left frontal, left temporal, and left parietal lobes, as shown in Figure 7. The results observed are consistent with previous research (Ernst et al. 2004; Ahmad et al. 2016; Sandor et al. 2018; Guidotti et al. 2019). The F3 electrode in the frontal lobe did not record clear negative waves and was excluded. Therefore, the F7 and FC5 electrodes in the frontal lobe, T7 and P7 electrodes in the temporal lobe, and CP5 and P3 electrodes in the parietal lobe were selected as analysis electrodes. Additionally, segments within the range of $0 \sim 200$ ms before the key press were selected for further analysis. The waveform diagram of the six electrodes is shown in Figure 8.

3.2.1. Analysis of ERP amplitude

The average amplitude recorded by each electrode in the predetermined time window (within 200 ms before the key press) according to varying levels of interface complexity is shown in Table 3.

A 3×6 repeated-measures ANOVA on interface complexity (low, moderate, and high) and electrode (F7, FC5, T7, P7, CP5, and P3) was performed. Mauchly's test of sphericity showed that the electrode data violated the assumption of sphericity ($p < 0.05$). To correct for the degrees of freedom, Greenhouse–Geisser estimates of sphericity were adopted. The main effect of electrode was significant, F (1.839, 97.485) $=$ 21.821, $p < 0.001$, $\eta_p{}^2 = 0.292$. Mauchly's test of sphericity indicated that the interface complexity data met the assumption of sphericity ($p > 0.05$). The main effect of interface complexity was significant, F (2, 106) $=$ 3.153, $p = 0.047 < 0.05$, $\eta_p{}^2 = 0.056$. Mauchly's test of sphericity showed that the interaction between interface complexity and electrode violated the assumption of sphericity ($p < 0.05$). To correct for the degrees of freedom, Greenhouse–Geisser estimates of sphericity were adopted. The main effect of interaction between interface complexity and electrode was not significant, F (4.427, 234.639) $=$ 1.444, $p = 0.216 > 0.05$, $\eta_p{}^2 = 0.027$.

Figure 9 shows that the amplitudes of the three interface complexity levels were relatively comparable at the F7 and FC5 electrodes, while there were amplitude differences at the T7, P7, CP5 and P3 electrodes. One-way

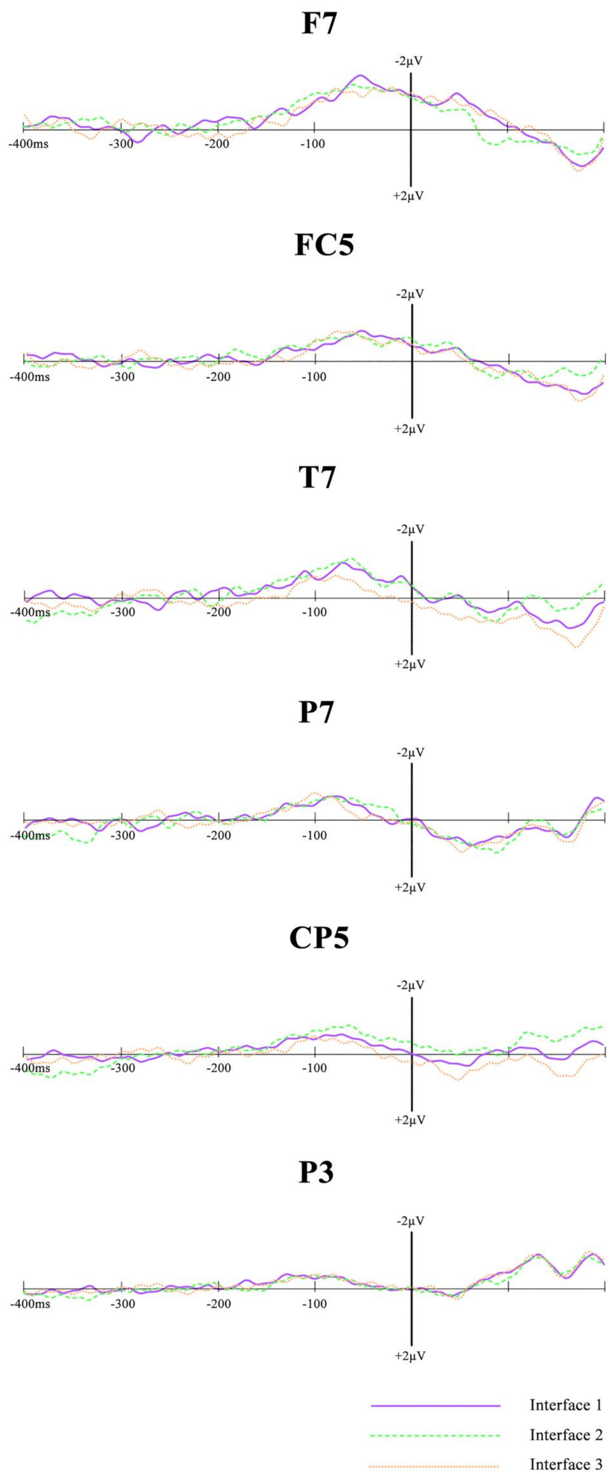


Figure 8. Waveforms at the F7, FC5, T7, P7, CP5, and P3 electrodes.

Table 3. Mean amplitude at each electrode (μV).

Electrode	Interface	Mean	Standard deviation	Participants (n)
F7	1	−2.5022	2.5361	54
	2	−2.6744	1.8144	54
	3	−2.4524	2.3559	54
FC5	1	−1.5317	1.3776	54
	2	−1.8220	1.2084	54
	3	−1.7930	1.5764	54
T7	1	−2.1919	1.4782	54
	2	−2.5902	1.5149	54
	3	−1.8933	1.2786	54
P7	1	−1.8559	1.1595	54
	2	−2.1381	1.1341	54
	3	−1.7841	1.1109	54
CP5	1	−1.4356	1.0305	54
	2	−1.8709	1.6126	54
	3	−1.2976	0.9812	54
P3	1	−1.2063	0.5240	54
	2	−1.2885	0.5460	54
	3	−1.0167	0.6062	54

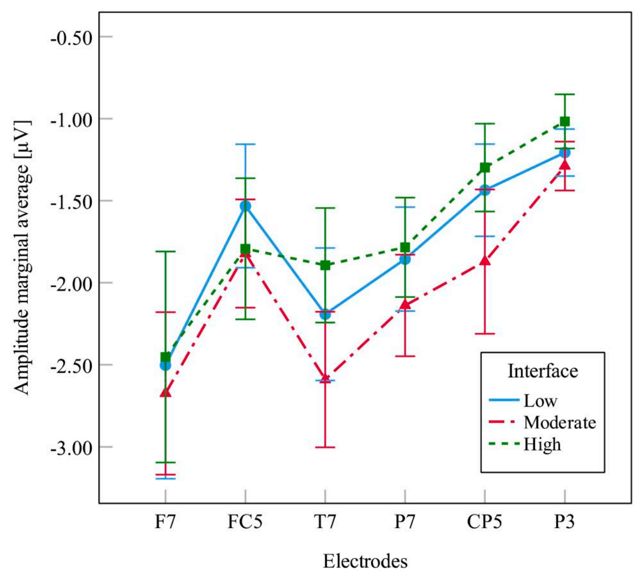


Figure 9. Estimated marginal average of the amplitude.

ANOVAs were performed to analyse the amplitudes at the T7, P7, CP5 and P3 electrodes. The main effect of interface complexity on amplitude at the T7 electrode was significant, F (2, 159) $=$ 3.238, $p = 0.042 < 0.05$, $\eta_p^2 = 0.039$. *Post hoc* analyses using Tukey's HSD indicated that amplitude at the T7 electrode differed significantly between moderate and high complexity ($p = 0.032$) but did not differ significantly between moderate and low complexity ($p = 0.318$). The main effect of interface complexity on amplitude at the P7 electrode was not significant, F (2, 159) $=$ 1.468, $p = 0.233 > 0.05$, $\eta_p^2 = 0.018$. Additionally, the main effect of interface complexity on amplitude at the CP5 electrode was significant, F (2, 159) $=$ 3.136, $p = 0.046 < 0.05$, $\eta_p^2 = 0.038$. *Post hoc* analyses using Tukey's HSD indicated that amplitude at the CP5 electrode differed significantly between moderate and high complexity ($p = 0.046$) but did not differ significantly between moderate and low complexity ($p = 0.166$). Furthermore, the main effect of interface complexity on amplitude at the P3 electrode was significant, F (2, 159) $=$ 3.349, $p = 0.038 < 0.05$, $\eta_p^2 = 0.040$. *Post hoc* analyses using Tukey's HSD indicated that amplitude at the P3 electrode differed significantly between

Table 4. Mean latency at each electrode (ms).

Electrode	Interface	Mean	Standard deviation	Participants (*n*)
F7	1	−74.70	50.76	54
	2	−88.37	55.99	54
	3	−69.33	45.51	54
FC5	1	−75.33	51.79	54
	2	−77.26	51.64	54
	3	−73.48	49.90	54
T7	1	−80.00	51.35	54
	2	−94.76	42.38	54
	3	−84.41	52.77	54
P7	1	−94.26	39.83	54
	2	−81.85	39.50	54
	3	−102.28	42.50	54
CP5	1	−79.57	45.48	54
	2	−72.93	45.91	54
	3	−93.78	41.61	54
P3	1	−105.46	52.76	54
	2	−87.09	44.25	54
	3	−108.09	42.54	54

moderate and high complexity ($p = 0.034$) but did not differ significantly between moderate and low complexity ($p = 0.726$).

3.2.2. Analysis of ERP latency

Since the ERP components of the experiment were response-locked and decision-making occurred before the key press reference point, all latencies were negative (Roggeveen, Prime, and Ward 2007). These latencies referred to the duration before the key press. In 200-ms time window before the key press, the average latency was recorded by each electrode according to interface complexity; these data are shown in Table 4.

A 3×6 repeated-measures ANOVA on interface complexity (low, moderate, and high) and electrode (F7, FC5, T7, P7, CP5, and P3) was performed. Mauchly's test of sphericity showed that the electrode data violated the assumption of sphericity ($p < 0.05$). To correct for the degrees of freedom, Greenhouse–Geisser estimates of sphericity were adopted. The main effect of electrode was significant, F (3.662, 194.096) = 7.229, $p < 0.001$, η_p^2 = 0.120. Mauchly's test of sphericity indicated that the interface complexity data met the assumption of sphericity ($p > 0.05$). The main effect of interface complexity was not significant, F (2, 106) = 0.411, $p = 0.664 > 0.05$, η_p^2 = 0.008. Mauchly's test of sphericity showed that the interaction between interface complexity and electrode violated the assumption of sphericity ($p < 0.05$). To correct for the degrees of freedom, Greenhouse–Geisser estimates of sphericity were adopted. The main effect of interaction between interface complexity and electrode was significant, F (7.765, 411.529) = 3.696, $p < 0.001$, $\eta_p^2 = 0.065$.

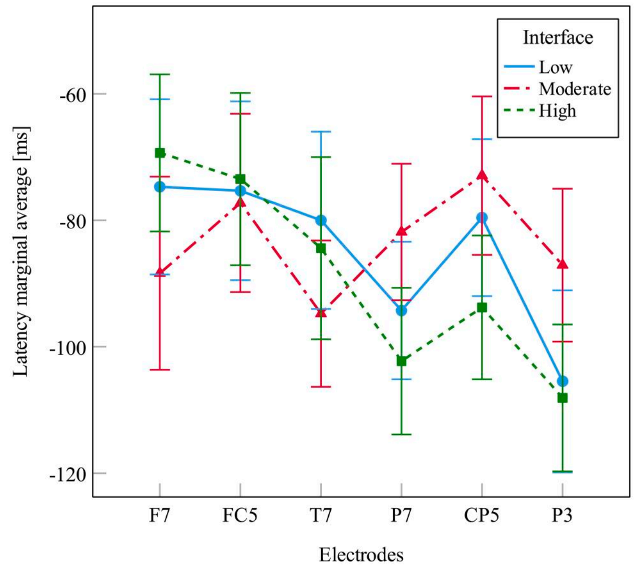


Figure 10. Estimated marginal average of the latency.

Figure 10 shows that the latencies of the three interface complexity levels were relatively comparable at the FC5 electrode, while there were latency differences at the F7, T7, P7, CP5 and P3 electrodes. One-way ANOVAs were performed to analyse the latencies at the F7, T7, P7, CP5 and P3 electrodes. The main effect of interface complexity on latency at the F7 electrode was not significant, F (2, 159) = 2.005, $p = 0.138 > 0.05$, η_p^2 = 0.025. The main effect of interface complexity on latency at the T7 electrode was not significant, F (2, 159) = 1.288, $p = 0.279 > 0.05$, η_p^2 = 0.016. The main effect of interface complexity on latency at the P7 electrode was significant, F (2, 159) = 3.464, $p = 0.034 < 0.05$, η_p^2 = 0.042. *Post hoc* analyses using Tukey's HSD indicated that latency at the P7 electrode differed significantly between moderate and high complexity ($p = 0.027$) but did not differ significantly between moderate and low complexity ($p = 0.254$). Additionally, the main effect of interface complexity on latency at the CP5 electrode was also significant, F (2, 159) = 3.111, $p = 0.047 < 0.05$, η_p^2 = 0.038. *Post hoc* analyses using Tukey's HSD indicated that latency at the CP5 electrode differed significantly between moderate and high complexity ($p = 0.041$) but did not differ significantly between moderate and low complexity ($p = 0.717$). Furthermore, the main effect of interface complexity on latency at the P3 electrode was significant, F (2, 159) = 3.236, $p = 0.042 < 0.05$, η_p^2 = 0.039. *Post hoc* analyses using Tukey's HSD indicated that latency at the P3 electrode differed significantly between moderate and high complexity ($p = 0.054$) but did not differ significantly between moderate and low complexity ($p = 0.105$).

4. Discussion

The behavioural results showed that the ACC and RT significantly differed according to interface complexity. The greater the interface complexity was, the lower the ACC and the longer the RT. Jin et al. (2017) indicated that a user's RT would increase as the amount of information increased. This is mainly because human cognitive resources are limited (Smalt et al. 2020). When users need to process more information, there is a shortage of cognitive resources, resulting in a lower ACC and longer duration needed to reach a decision. The results also showed that the ACC of trained participants could be maintained above 96% as the amount of information increased but the RT increased significantly. This indicates that excessive information hinders the effective decision-making of well-trained participants; that is, there is an upper limit of image entropy (a measure of interface complexity) suitable for decision-making.

The behavioural results also showed that for low interface complexity, the ACC was close to 100% and the RT was less than 2,000 ms. This indicates that little information does not hinder the effective decision-making of well-trained participants, or at least lack of evidence that there is a lower limit of image entropy suitable for decision-making. Therefore, the existence of a lower limit needs to be confirmed by ERP experiments with high temporal resolution.

The ERP source analysis showed that brain activation during user decision-making was mainly located in the left hemisphere, which is consistent with previous research. Euston, Gruber, and Mcnaughton (2012) indicated that the medial prefrontal lobe is activated in decision-making. Ungerleider, Courtney, and Haxby (1998) and Gitelman et al. (1999) suggested that the dorsal anterior cingulate maps to conflict monitoring and that the left premotor area maps to movement preparation. These relationships were also consistent with the trend of left lateralisation for mathematical tasks (Burbaud et al. 1999). Differing interface complexity produced obvious amplitude changes 100 ms before the key press. Kitajima and Toyota (2013) believed that the cognitive band appears approximately 100 ms before an action. The operator's decision-making task is divided into three parts: monitoring, decision-making and action execution (Liu et al. 2020). The decision-making process is completed before the key is pressed; thus, the potential detected approximately 100 ms before the key press may be related to the decision-making process. The decision-making process can also be divided into early information processing and subsequent action selection (Maksimenko et al. 2020). The temporal lobe is related to the former aspect of the process, the frontal lobe is related to the latter, and the parietal lobe is related to both (Horr, Braun, and Volz 2014; Guidotti et al. 2019; Zhou and Freedman 2019; Maksimenko et al. 2020). Consistent with these findings, the T7, P7, CP5, and P3 electrodes in the present study were sensitive to changes in interface complexity.

The amplitude of each interface complexity level significantly differed at the T7, CP5 and P3 electrodes. According to post hoc analysis, the amplitude of moderate complexity was significantly higher than that of high complexity (in absolute value), while the amplitude of low complexity was generally in between. The same order was observed at the P7 electrode. The F7 and FC5 electrodes were not sensitive to changes in interface complexity, and the amplitudes at these electrodes for each interface complexity level were relatively approximate. The amplitude at the T7, P7, CP5 and P3 electrodes for each interface complexity level showed that the amplitude was highest in response to moderate complexity. Higher amplitude indicates better user decision-making, i.e. more attention, deeper thinking, and more decisive action (Feng et al. 2010, Garrido-Chaves et al. 2021). Low complexity, with the lowest complexity, did not elicit the highest ERP amplitude, implying that less interface information does not lead to better decision-making; that is, there is a lower limit of image entropy (or interface complexity) suitable for decision-making.

The latency of each interface complexity level significantly differed at the P7, CP5 and P3 electrodes. According to post hoc analysis, the latency of moderate complexity was significantly shorter than that of high complexity (in absolute value), while the latency of low complexity was generally in between. In the frontal lobe (the F7 and FC5 electrodes), no obvious difference was observed. Euston, Gruber, and Mcnaughton (2012) suggested that the activation of the frontal lobe is related to the ability to extract memories of corresponding information, which represents the late stage of decision-making. Therefore, frontal lobe electrode data were not used as the basis for analysis. Low complexity, which provided the least amount of information, did not have the shortest latency, implying a lower limit of image entropy (or interface complexity) suitable for decision-making.

The above analyses demonstrate that interface complexity has upper and lower limits suitable for decision-making. This enables a new method of interface evaluation that uses image entropy as the indicator and the upper and lower limits of interface

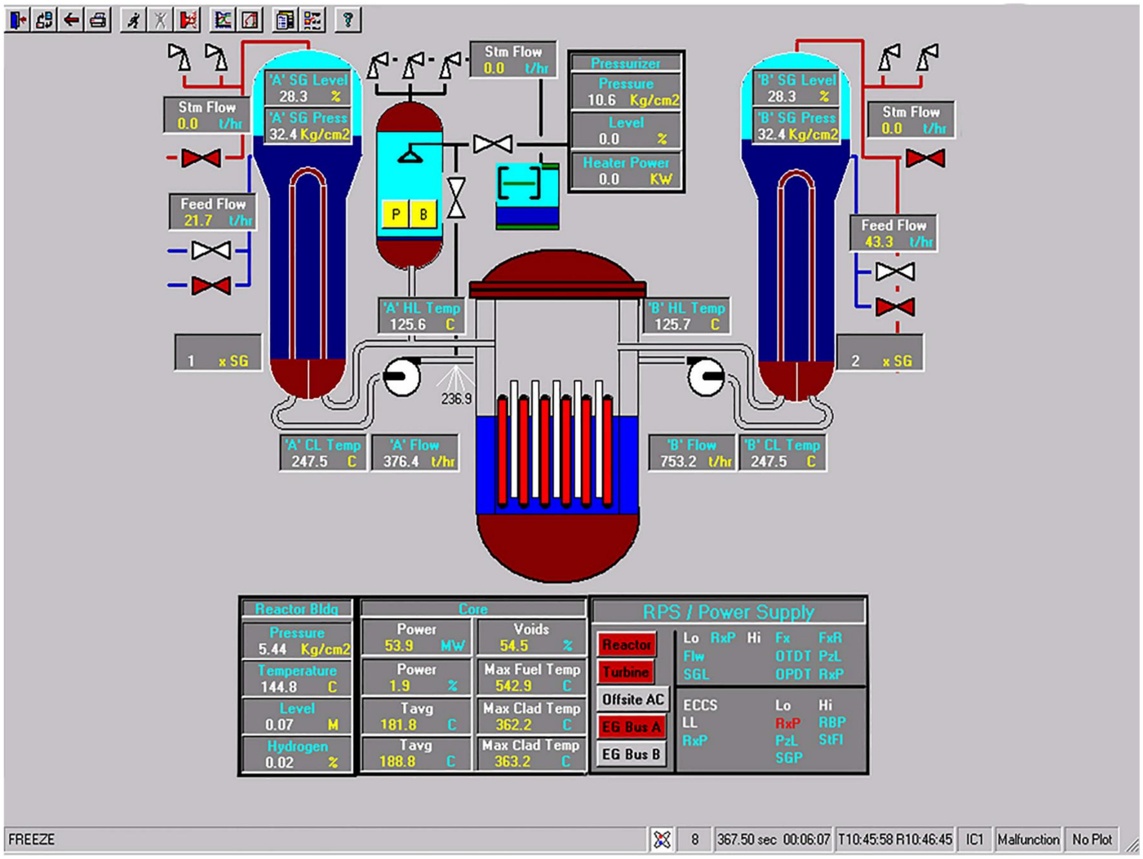


Figure 11. NPP main display interface.

complexity as the threshold. This method thus evaluates an interface from the perspective of its benefits to user decision-making and is fast, economical, and suitable for all design phases of NPP interfaces.

The general procedure of the evaluation method is as follows. First, the interface should be converted into an 8-bit grey image. Second, the probability distribution of the grey levels in the image and the image entropy should be calculated. Third, the calculation results should be compared with the threshold of this type of interface to draw a conclusion. If the values fall within the threshold, then the interface should be considered good, and if the values fall outside the threshold, then the interface should be considered not good. For interfaces that fail the evaluation, follow-up suggestions for improvement are as follows: if the entropy value is lower than the threshold value, then the interface complexity needs to be increased, and if the entropy value is higher than the threshold value, then the interface complexity needs to be reduced.

Interface thresholds can be determined by using questionnaire or experimental methods. The questionnaire method requires participants to complete a scale that includes images of interfaces with various entropy levels. Participants score the questionnaire based on their own experience, and the scores are used to derive the upper and lower thresholds of the interface. The experimental method requires participants to complete a set of interface decision-making experiments. The experimental stimuli consist of interfaces presenting various amounts of information. The performance results are used to derive the upper and lower thresholds of the interface.

For example, the NPP main display interface can be evaluated. A main display interface of the generic two-loop pressurised water reactor was designed, as shown in Figure 11 (at a resolution of 1024×768 pixels).

First, a screenshot of the interface was created, and the picture was converted to an 8-bit greyscale image. Second, the image entropy function was applied to calculate the image entropy, and the result was 1.7232. This value is within the interface threshold of 1.28–1.79 (the threshold obtained by the questionnaire method, see Appendix); therefore, the design of the interface was good.

The evaluation time for a single interface was less than one minute, and the cost was almost zero. This method is suitable not only for the rapid evaluation of a large number of interfaces but also for interfaces at all phases of design for real-time adjustment.

The disadvantage of this method is that the threshold value needs to be clarified before evaluation; this value changes according to design standards and tasks. Fortunately, the threshold value is easy to determine because NPP interfaces have specific design standards and tasks. Future research should optimise the threshold determination process and construct a comprehensive evaluation system along with existing evaluation methods, such as heuristic evaluation, Hick's law and experimental evaluation, to comprehensively evaluate NPP interfaces.

5. Conclusions

The aim of this study was to understand the relationship between interface complexity and user decision-making by conducting ERP experiments. Three NPP interfaces with gradually increasing image entropy (interface complexity) were used as stimuli. The decision-making task required the user to press different keys based on information provided by the interface. Six electrodes, namely, the F7, FC5, T7, P7, CP5, and P3 electrodes, exhibited obvious amplitude changes during decision-making.

This study provided information that deepened the current understanding of decision-making from the perspective of interface complexity. The main findings are as follows: there is a specific range of interface complexity (image entropy) that facilitates decision-making. The performance results and cognitive resource theory indicate the upper limit of image entropy (interface complexity), and the ERP results, indicate the lower limit of image entropy (interface complexity).

This study also proposed a novel method of interface evaluation based on the influence of interface complexity on user decision-making. The method can be used to determine whether an interface is conducive to decision-making based on whether the interface image entropy (interface complexity) falls within a specific range. The method flow is as follows: first, convert interfaces into grey images, then calculate the probability distribution of grey levels and the image entropy and, finally, compare the image entropy with the recommended threshold. Compared with heuristic evaluation, this method has the advantages of objectivity and consistency. Compared with Hick's Law, this method is more suitable for complex situations that require substantial reading and careful consideration. Compared with experimental evaluation, this method has the advantages of being fast and low-cost. Moreover, this method solves the problem of evaluating interfaces in the initial design phase.

The shortcoming of this study is that although interface complexity is thought to exhibit the most appropriate image entropy range for decision-making, the combined influence of other design factors, such as colour, layout, font type and aesthetics (Tuch et al. 2012; Lazard and King 2020), on decision-making was not considered. Aesthetics are important when designing digital interfaces for complex information systems. An elegant interface with a usable layout may exert different influences on user decision-making than an ugly interface with a cluttered layout, even if the two interfaces have the same level of complexity. Therefore, future research should focus on the combined influence of interface design elements, such as interface complexity, colour, layout, font type and aesthetics, on decision-making to fully comprehend the influence of an interface on user decision-making.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Natural Science Foundation of China under grant [No. 71871056, 71801037, 71901061, and 72171044], the National Aerospace Science Foundation of China under grant [No. 20200058069002], and ZhiShan Scholar Program of Southeast University (2242022R40004)

References

Ahmad, F., A. Hussain, S.U. Chaudhary, I. Ahmad, and S.M. Ramay. 2016. "A Novel Method for Detection of Voxels for Decision Making: An Fmri Study." *International Journal of Imaging Systems and Technology* 26 (2): 163–167. doi:10.1002/ima.22172.

Ba, Y., W. Zhang, Q. Peng, G. Salvendy, and D. Crundall. 2016. "Risk-Taking on the Road and in the Mind: Behavioural and Neural Patterns of Decision Making between Risky and Safe Drivers." *Ergonomics* 59 (1): 27–38. doi:10.1080/00140139.2015.1056236.

Bohanec, M., I. Vrbanic, I. Basic, K. Debelak, and L. Strubelj. 2020. "A Decision-Support Approach to Severe Accident Management in Nuclear Power Plants." *Journal of Decision Systems* 29 (sup1): 438–449. doi:10.1080/12460125.2020.1854426.

Burbaud, P., O. Camus, D. Guehl, B. Bioulac, J.M. Caille, and M. Allard. 1999. "A Functional Magnetic Resonance Imaging Study of Mental Subtraction in Human Subjects." *Neuroscience Letters* 273 (3): 195–199. doi:10.1016/S0304-3940(99)00641-2.

Changoluisa, V., P. Varona, and F. De Borja Rodriguez. 2020. "A Low-Cost Computational Method for Characterizing Event-Related Potentials for Bci Applications and beyond."

IEEE Access.8: 111089–111101. doi:10.1109/ACCESS.2020.3000187.

Chen, H., L. Pang, X. Wanyan, S. Liu, Y. Fang, and D. Tao. 2021. "Effects of Air Route Alternation and Display Design on an Operator's Situation Awareness, Task Performance and Mental Workload in Simulated Flight Tasks." *Applied Sciences* 11 (12): 5745. doi:10.3390/app11125745.

Chen, K., Z. Li, and G.A. Jamieson. 2018. "Influence of Information Layout on Diagnosis Performance." *IEEE Transactions on Human-Machine Systems* 48 (3): 316–323. doi:10.1109/THMS.2017.2767284.

Cheng, F.-F., C.-S. Wu, and B. Leiner. 2019. "The Influence of User Interface Design on Consumer Perceptions: A Cross-Cultural Comparison." *Computers in Human Behavior* 101: 394–401. doi:10.1016/j.chb.2018.08.015.

Collins, D.J., A. Abbott, and H. Richards. 2011. *Performance Psychology e-Book: A Practitioner's Guide*. Amsterdam: Elsevier Health Sciences.

Diaz-Lago, M, and H. Matute. 2019. "A Hard to Read Font Reduces the Causality Bias." *Judgment and Decision Making* 14 (5): 547–554.

Ernst, M., E.E. Nelson, E.B. Mcclure, C.S. Monk, S. Munson, N. Eshel, E. Zarahn, E. Leibenluft, A. Zametkin, K. Towbin, J. Blair, D. Charney, and D.S. Pine. 2004. "Choice Selection and Reward Anticipation: An Fmri Study." *Neuropsychologia* 42 (12): 1585–1597. doi:10.1016/j.neuropsychologia.2004.05.011.

Euston, D.R., A.J. Gruber, and B.L. Mcnaughton. 2012. "The Role of Medial Prefrontal Cortex in Memory and Decision Making." *Neuron* 76 (6): 1057–1070. 10.1016/j.neuron.2012.12.002

Feng, Y., C.-L. Zhou, J.-C. Zhang, and M.-L. Tian. 2010. "Neural Mechanisms of Intuitive Tactical Decision-Making Predominance of High-Level Fencing Athletes." *Journal of Medical and Biological Engineering* 30 (1): 47–56.

Garrido-Chaves, R., V. Perez, M. Perez-Alarcon, I. Crespo-Sanmiguel, T.O. Paiva, V. Hidalgo, M.M. Pulopulos, and A. Salvador. 2021. "Subjective Memory Complaints and Decision Making in Young and Older Adults: An Event-Related Potential Study." *Frontiers in Aging Neuroscience* 13: 1–26. doi:10.3389/fnagi.2021.695275.

Gitelman, D.R., A.C. Nobre, T.B. Parrish, K.S. Labar, Y.H. Kim, J.R. Meyer, and M.M. Mesulam. 1999. "A Large-Scale Distributed Network for Covert Spatial Attention – Further Anatomical Delineation Based on Stringent Behavioural and Cognitive Controls." *Brain* 122 (6): 1093–1106. doi:10.1093/brain/122.6.1093.

Guidotti, R., A. Tosoni, M.G. Perrucci, and C. Sestieri. 2019. "Choice-Predictive Activity in Parietal Cortex during Source Memory Decisions." *NeuroImage* 189: 589–600. doi:10.1016/j.neuroimage.2019.01.071.

Guo, F., J. Chen, M. Li, W. Lyu, and J. Zhang. 2021. "Effects of Visual Complexity on User Search Behavior and Satisfaction: An Eye-Tracking Study of Mobile News Apps." *Universal Access in the Information Society* : 1–14. doi:10.1007/s10209-021-00815-1.

Han, X., Q. Xue, and Q. Zhang. 2017. "Information Providing in a Task-Centric Military in-Vehicle Adaptive Human-Computer Interface Based on Operator Decision-Making Characteristic." *Advances in Mechanical Engineering* 9 (6): 168781401770993. doi:10.1177/1687814017709935.

Horr, N.K., C. Braun, and K.G. Volz. 2014. "Feeling before Knowing Why: The Role of the Orbitofrontal Cortex in Intuitive Judgments-An Meg Study." *Cognitive, Affective & Behavioral Neuroscience* 14 (4): 1271–1285. doi:10.3758/s13415-014-0286-7.

Hugo, J.V, and D.I. Gertman. 2016. "A Method to Select Human-System Interfaces for Nuclear Power Plants." *Nuclear Engineering and Technology* 48 (1): 87–97. doi:10.1016/j.net.2015.10.004.

Jin, L., C. Hong-Zhen, Q. Yue, and L. Xun. 2017. "Can "Stay Online" Reduce Our Work Performance?" In 2017 24th Annual International Conference on Management Science & Engineering, edited by M.A. Tao, 623–630. Nomi, Japan: Institute of Electrical and Electronics Engineers (IEEE).

Kitajima, M, and M. Toyota. 2013. "Decision-Making and Action Selection in Two Minds: An Analysis Based on Model Human Processor with Realtime Constraints (Mhp/rt)." *Biologically Inspired Cognitive Architectures* 5: 82–93. doi:10.1016/j.bica.2013.05.003.

Korn, C.W., J. Ries, L. Schalk, Y. Oganian, and H. Saalbach. 2018. "A Hard-to-Read Font Reduces the Framing Effect in a Large Sample." *Psychonomic Bulletin & Review* 25 (2): 696–703. doi:10.3758/s13423-017-1395-4.

Lazard, A.J, and A.J. King. 2020. "Objective Design to Subjective Evaluations: Connecting Visual Complexity to Aesthetic and Usability Assessments of Ehealth." *International Journal of Human–Computer Interaction* 36 (1): 95–104. doi:10.1080/10447318.2019.1606976.

Li, J., Z.L. Yu, Z. Gu, W. Wu, Y. Li, and L. Jin. 2018. "A Hybrid Network for Erp Detection and Analysis Based on Restricted Boltzmann Machine." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26 (3): 563–572. doi:10.1109/TNSRE.2018.2803066.

Liu, Y., Y. Luan, G. Zhang, H. Hu, J. Jiang, L. Zhang, T. Qing, Y. Zou, D. Yang, and L. Xi. 2020. "Human Reliability Analysis for Operators in the Digital Main Control Rooms of Nuclear Power Plants." *Journal of Nuclear Science and Technology* 57 (6): 719–733. doi:10.1080/00223131.2020.1720848.

Maglic, L, and D. Zec. 2020. "The Impact of Bridge Alerts on Navigating Officers." *Journal of Navigation* 73 (2): 421–432. doi:10.1017/S0373463319000687.

Maksimenko, V.A., A. Kuc, N.S. Frolov, M.V. Khramova, A.N. Pisarchik, and A.E. Hramov. 2020. "Dissociating Cognitive Processes During Ambiguous Information Processing in Perceptual Decision-Making." *Frontiers in Behavioral Neuroscience* 14: 95. doi:10.3389/fnbeh.2020.00095.

Meng, J, and G. Xiu. 2018. "Objective Decision-Making Brain Mechanism of Public- Private-Partnerships Project Risk Management Based on Decision Neuroscience Theory." *NeuroQuantology* 16 (5): 284–289. doi:10.14704/nq.2018.16.5.1249.

Oyibo, K, and J. Vassileva. 2020. "The Effect of Layout and Colour Temperature on the Perception of Tourism Websites for Mobile Devices." *Multimodal Technologies and Interaction* 4 (1): 8. doi:10.3390/mti4010008.

Petrovcic, A., M. Setinc, T. Burnik, and V. Dolnicar. 2018. "A Comparison of the Usability of a Standard and an Age-Friendly Smartphone Launcher: Experimental Evidence from Usability Testing with Older Adults." *Internationale Zeitschrift Fur Rehabilitationsforschung. Revue Internationale de Recherches de Readaptation* [International Journal of

Rehabilitation Research] 41 (4): 337–342. doi:10.1097/MRR.0000000000000307.

Roggeveen, A.B., D.J. Prime, and L.M. Ward. 2007. "Lateralized Readiness Potentials Reveal Motor Slowing in the Aging Brain." *The Journals of Gerontology-Series B, Psychological Sciences and Social Sciences* 62 (2): P78–P84. doi:10.1093/geronb/62.2.p78.

Sandor, S., Ş. Delil, S. Yağcı, B. Korkmaz, and S.N. Yeni. 2018. "Improved Decision-Making and Psychophysiological Responses in Mesial Temporal Lobe Epilepsy after Anterior Temporal Lobectomy." *Epileptic Disorders* 20 (6): 517–524. doi:10.1684/epd.2018.1019.

Shen, L, and O. Urminsky. 2013. "Making Sense of Nonsense: The Visual Salience of Units Determines Sensitivity to Magnitude." *Psychological Science* 24 (3): 297–304. doi:10.1177/0956797612451470.

Shukla, V., G. Auriol, and K.W. Hipel. 2016. "Multicriteria Decision-Making Methodology for Systems Engineering." *IEEE Systems Journal* 10 (1): 4–14. doi:10.1109/JSYST.2014.2344635.

Smalt, C.J., P.T. Calamia, A.P. Dumas, J.P. Perricone, T. Patel, J. Bobrow, P.P. Collins, M.L. Markey, and T.F. Quatieri. 2020. "The Effect of Hearing-Protection Devices on Auditory Situational Awareness and Listening Effort." *Ear and Hearing* 41 (1): 82–94. doi:10.1097/AUD.0000000000000733.

Starke, S.D, and C. Baber. 2018. "The Effect of Four User Interface Concepts on Visual Scan Pattern Similarity and Information Foraging in a Complex Decision Making Task." *Applied Ergonomics* 70: 6–17. doi:10.1016/j.apergo.2018.01.010.

Tuch, A.N., E.E. Presslaber, M. Stöcklin, K. Opwis, and J.A. Bargas-Avila. 2012. "The Role of Visual Complexity and Prototypicality regarding First Impression of Websites: Working towards Understanding Aesthetic Judgments." *International Journal of Human-Computer Studies* 70 (11): 794–811. doi:10.1016/j.ijhcs.2012.06.003.

Ungerleider, L.G., S.M. Courtney, and J.V. Haxby. 1998. "A Neural System for Human Visual Working Memory." *Proceedings of the National Academy of Sciences of the United States of America* 95 (3): 883–890. doi:10.1073/pnas.95.3.883.

Van Zon, N.C.M., C. Borst, D.M. Pool, and M.M. Van Paassen. 2020. "Touchscreens for Aircraft Navigation Tasks: Comparing Accuracy and Throughput of Three Flight Deck Interfaces Using Fitts' Law." *Human Factors* 62 (6): 897–908. doi:10.1177/0018720819862146.

Vincent, K., R.E. Roth, S.A. Moore, Q. Huang, N. Lally, C.M. Sack, E. Nost, and H. Rosenfeld. 2019. "Improving Spatial Decision Making Using Interactive Maps: An Empirical Study on Interface Complexity and Decision Complexity in the North american Hazardous Waste Trade." *Environment and Planning B: Urban Analytics and City Science* 46 (9): 1706–1723. doi:10.1177/2399808318764122.

Westerman, S.J., E.J. Sutherland, P.H. Gardner, R. Metcalfe, J. Nash, S. Palframan, and N. Woodburn. 2012. "Ecommerce Interface Colour and Consumer Decision Making: Two Routes of Influence." *Color Research & Application* 37 (4): 292–301. doi:10.1002/col.20690.

Zhang, R. 2018. "Decision-Making Mechanism under Economic Management Risk Based on Event-Related Potential Neuroimaging Technique." *NeuroQuantology* 16 (3): 75–82. doi:10.14704/nq.2018.16.3.1202.

Zhou, Y, and D.J. Freedman. 2019. "Posterior Parietal Cortex Plays a Causal Role in Perceptual and Categorical Decisions." *Science* 365 (6449): 180–185. doi:10.1126/science.aaw8347.

Appendix

Determination of the interface threshold based on the questionnaire

1. Questionnaire design

A questionnaire for NPP main display interfaces was designed as follows.

(a)

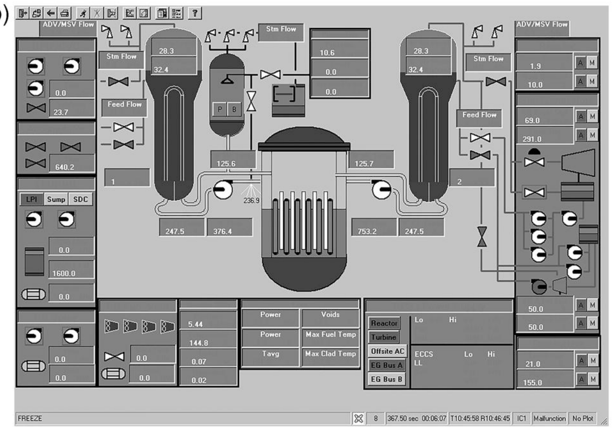
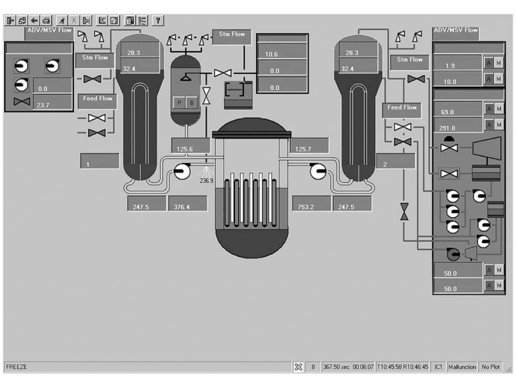
(b)


Figure 12. Reference interfaces.

Subjective evaluation scale

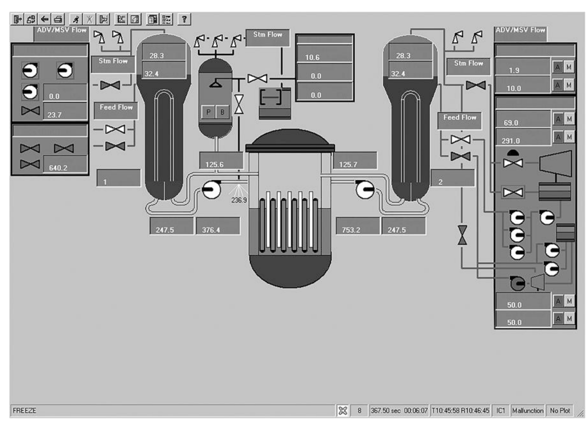
Interface complexity is divided into five categories: very low, low, suitable, high and very high. Please determine which of the five types of interface complexities is shown. To facilitate the evaluation of interface complexity, reference interfaces with a very low interface complexity and a very high interface complexity were provided, as shown in Figure 12.

No.1 Please rate the interface complexity.



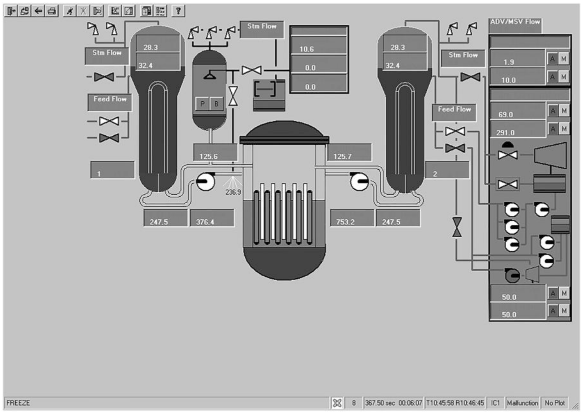
^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.2 Please rate the interface complexity.



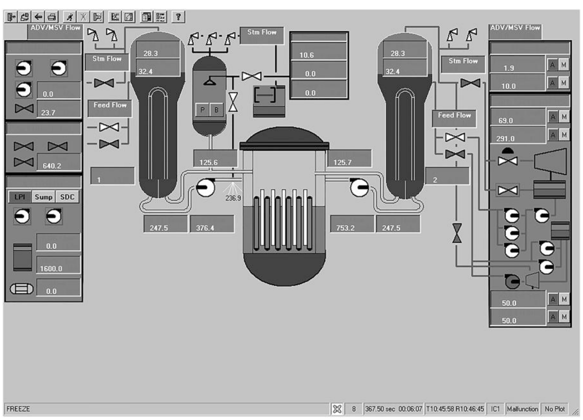
^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.3 Please rate the interface complexity.



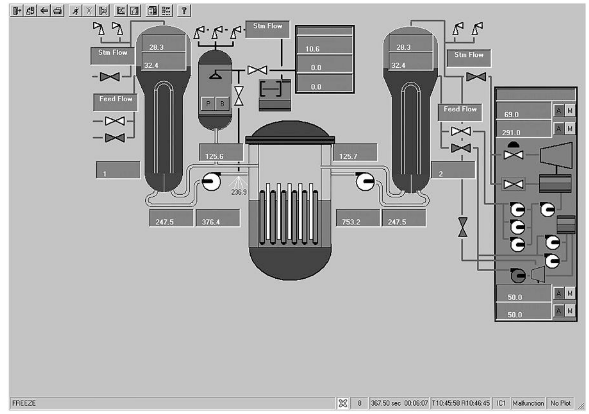
^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.4 Please rate the interface complexity.



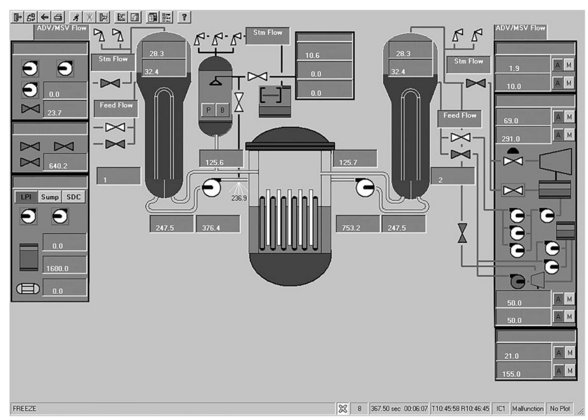
^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.5 Please rate the interface complexity.



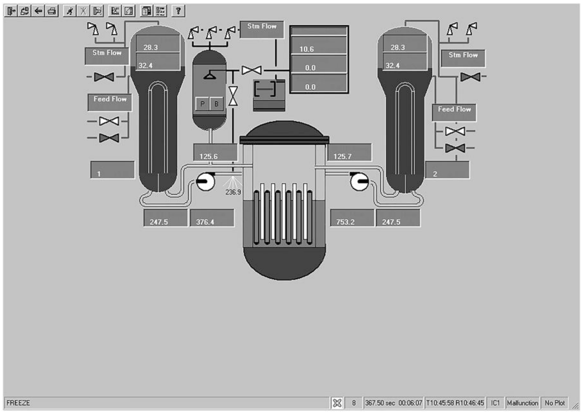
^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.6 Please rate the interface complexity.



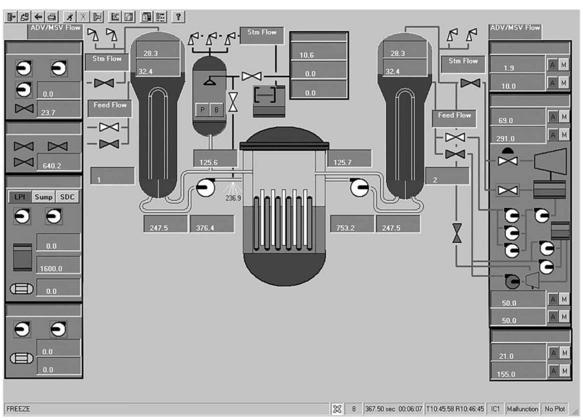
^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.7 Please rate the interface complexity.



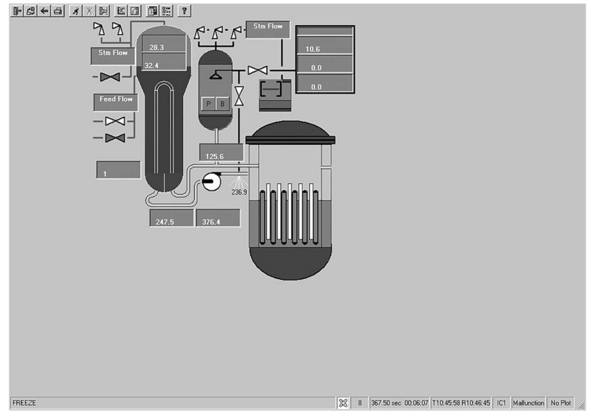
^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.8 Please rate the interface complexity.



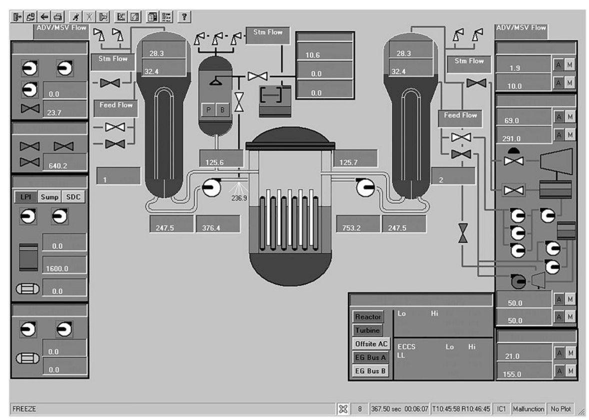
^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.9 Please rate the interface complexity.



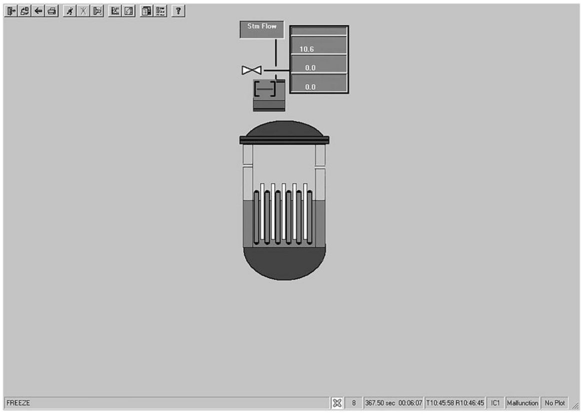
^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.10 Please rate the interface complexity.



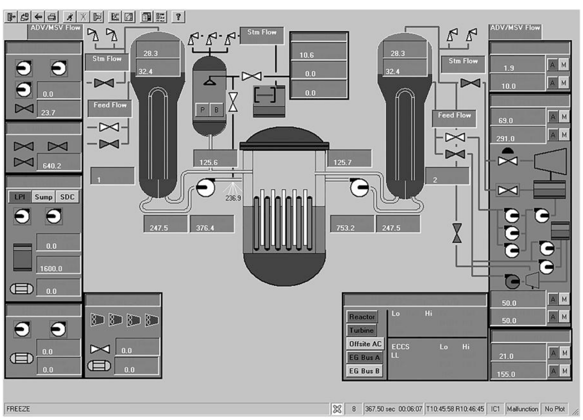
^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.11 Please rate the interface complexity.



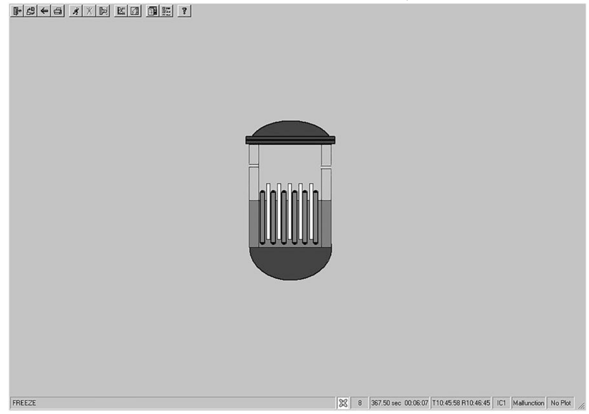
^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.12 Please rate the interface complexity.



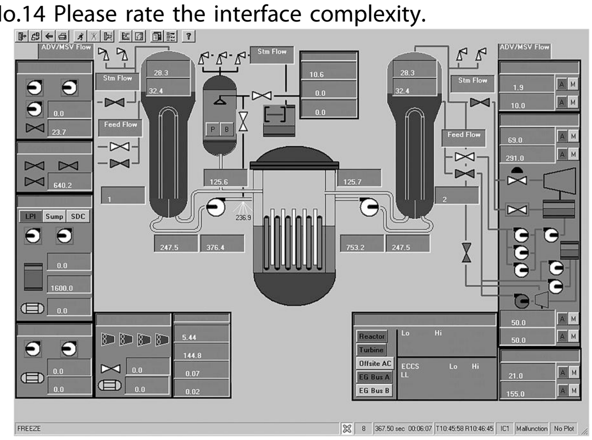
^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.13 Please rate the interface complexity.



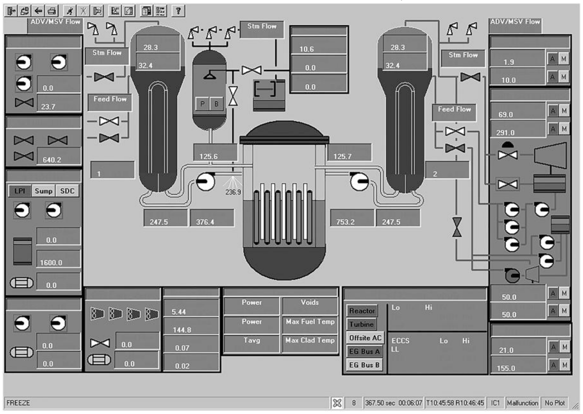
^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.14 Please rate the interface complexity.



^ Very low
^ Low
^ Suitable
^ High
^ Very high

No.15 Please rate the interface complexity.



^ Very low
^ Low
^ Suitable
^ High
^ Very high

2. Threshold determination

Questionnaire responses were obtained from 50 participants. These participants were all trained graduate students and had participated in many NPP interface design projects; thus, they were considered potential expert users.

Table 5. Questionnaire results.

Question number	Image entropy	Mode
1	1.7949	3
2	1.8434	2
3	1.6987	3
4	1.9532	2
5	1.6316	3
6	1.9961	2
7	1.2757	3
8	2.0505	2
9	1.1756	2
10	2.1466	1
11	0.5583	1
12	2.2005	1
13	0.3738	1
14	2.2308	1
15	2.2699	1

For each interface, complexity was ranked as very low (1 point), low (2 points), suitable (3 points), high (2 points), or very high (1 point). The interface with a mode of 3 points was deemed to be suitable for decision-making. The results of the questionnaire are shown in Table 5. The entropy range of the image with a mode of 3 points was 1.2757–1.7949 (at a resolution of 1024×768 pixels). The questionnaire data revealed that the empirical image entropy threshold for the NPP main system interface was 1.28–1.79.