Evaluation of Human-System Interfaces with Different Information Organization Using an Eye Tracker

Kejin Chen and Zhizhong Li*

Department of Industrial Engineering, Tsinghua University, Beijing, 100084, P.R. China
zzli@mail.tsinghua.edu.cn

Abstract. The increasing use of digitalized displays in the instrumentation & control systems of nuclear power plants has brought new issues related to human-computer interaction, especially under emergency circumstances that are known to be very stressful. This paper studies how interfaces with different information organization (functional layout vs. process layout) influence human-computer interaction behaviors as emergency occurs in terms of search efficiency, difficulty of information abstraction, and workload by using the eye tracking technique on a simulated platform. The result shows that the average blink rate and average blink numbers at the two levels of information organization were different significantly. This may indicate that the functional design was superior to the process design in user workload. The results did not prove the superiority of the functional interface design to the process one in search efficiency and difficulty of information abstraction, since no significant difference was found in the number of fixations and fixations duration mean values.

Keywords: Functional based task analysis, Information organization, Eye tracking.

1 Introduction

The Fukushima Daiichi nuclear power plant (NPP) Accident in 2011 raised again great public concern about the nuclear safety, after 2 decades since Chernobyl Accident. Tsunami and earthquake which directly resulted in Fukushima Daiichi nuclear accident is far beyond human's ability, while most accidents in safety-critical industries so far are caused by human errors. More than 90% accidents in NPPs, 80% ones in the chemical and petro-chemical industries, and over 70% of aviation accidents were caused by human errors (Adihikari et al., 2008).

It is therefore reasonable to focus on operators and their behaviors. In a NPP, the main control room (MCR), as a core part of the system, is expected to provide appropriate information and human-machine interface to facilitate operators' behaviors. The past decades have witnessed remarkable upgrades in MCRs, including its layout, devices and design concepts.

Conventional buttons, switches, and other analog equipment are gradually replaced, which is a prime improvement in MCRs. In NPPs of new generation like AP 1000, digital displays are extensively used. Such transform on the one hand caters to

* Corresponding author.

P.L.P. Rau (Ed.): CCD/HCII 2013, Part I, LNCS 8023, pp. 288–295, 2013.

the trend of technology advance, improving the safety and reliability of the systems; on the other hand, it brings high information burden, tedious interface management tasks, and other new problems (Kim & Seong, 2009). For this reason, designing a good digital interface in MCRs, which meets the task demand and the physical and psychological characteristics of operators, is of necessity. Several design philosophies have been put forward, for example, ecological interface design (EID) and 'User-centered design'. EID is based on two conceptual tools: the abstraction hierarchy and the skills, rules, and knowledge (SRK) taxonomy (Rasmussen, 1985). Review by Vicente (2002) can be referred to for specific instruction of EID. In the user-centered approach, task analysis is a key tool used widely to analyze user needs. What distinguishes EID from the user-centered approach lies mainly in the difference between abstraction hierarchy and task analysis.

The Function-Based Task Analysis (FBTA) proprietary to AP 1000 NPPs is one kind of task analysis. It breaks down goals in different levels according to their corresponding functions, providing which task information should be presented in the NPP interfaces. Wu et al.(2012) introduced FBTA in detail. Besides, in Wu's (2012) study, the functional display based on FBTA seemed better help understand system's operating status, ease the workload, while process displays helped the participants understand NPP's structure and working principles better. However, Wu's functional displays differ from the process displays principally in three aspects: information organization, information presentation and component representation. No research so far demonstrates which aspect influences the interaction and that to what extent the influence can be. This study aims to explore how interfaces with different information organization could influence human-computer interaction behaviors in terms of search efficiency, difficulty of information abstraction, and workload.

When users interact with an interface for a specific goal, they observe the interface, and articulate the desired task goal into input language. Since the observation and articulation are both cognitive behaviors, performance in the two processes is hardly measured directly. This study used eye tracking techniques to evaluate search efficiency, difficulty of information abstraction, and workload, which are supposed to reflect human cognition behaviors (Richardson & Spivey, 2004).

It is hypothesized in this study that the functional design is superior to the process design. Specifically, the functional design is supposed to possess higher search efficiency, less difficulty of information extraction, and less workload than the process one.

In the rest of this paper, Section 2 introduces the methodology of study, and Section 3 presents the results. Discussion and conclusions are presented separately in Sections 4 and 5.

2 Methodology

2.1 Independent Variables

The independent variable, information organization, has two levels: functional interface and process interface. Figure 1 illustrates the functional interface based on FBTA, and Figure 2 illustrates the process interface.

Fig. 1. Illustration of the functional interface

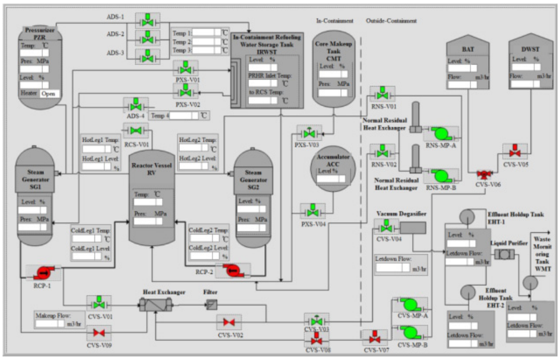

Fig. 2. Illustration of the process interface

There are two display screens. The left one is the main display for monitoring system status using the interfaces showed in Figures 1 and 2. The right one shows the computerized procedure for the participants to perform the emergency operations, as Figure 3 shows. The participants monitored the display on the left side, and took actions once the accident signal came up following the emergency operation procedure (EOP) presented on the right screen. The experiment was designed as within-subjects. When the participants finished the task with the functional (or process) interface, they took a short rest, and then began the task with process (or functional) interface. The order to show the interfaces was randomly arranged to reduce the learning effect. For each interface, the participants performed the EOP task 16 times. There were eight procedure routes in total and the participants should go through each route twice with every interface. The order of the routes was also randomly arranged by presenting the corresponding system status values (accident scenarios). Under such a circumstance, the participants had to carefully check system parameters from the presented interface to determine the selection of routes in the EOP. In this experiment, the initiative event was the High Pressurizer Level.

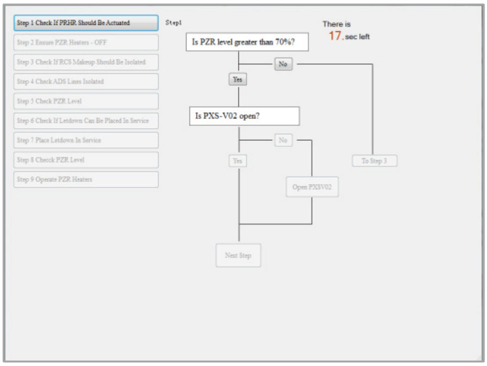

Fig. 3. Illustration of interface of the emergency operating procedure

2.2 Dependent Variables

When executing the tested EOP, the participants needed to observe the interface, collected information from it, and then followed branches of the EOP flow chart (as illustrated in Figure 3) in accordance with the current system status. Search efficiency and difficulty of information abstraction were the two aspects to evaluate whether the participants observed and collected information quickly or not. Number of fixations (Goldberg & Kotval, 1998) and fixation duration mean (Goldberg & Kotval, 1998) were measured for evaluating the two aspects. For each route in the EOP, the workload was assessed by blink rate and pupillary response to study whether the two interfaces required distinct cognitive effort.

All eye tracking measures in this study were recorded by the head-mounted eye tracker iView X from SMI© Company, with 50Hz sampling frequency.

Blink Rate: It refers to average number of blinking per minute. Brookling, Wilson and Swan (1996) reported that blink rate is more sensitive to workload than many other eye tracking measures. In this study, average blink rate for each interface was calculated.

Average Percentage Change of Pupil Size (APCPS): The participants were asked to relax their eyes before the experiment task was started, and their pupil sizes at the relaxed condition were recorded as the baseline. In this study, **APCPS** was calculated by formula (1), (2) and (3) (Iqbal et al., 2004).

$$Pupil\,size = \frac{Avg.Pupil\,Size\,X \times Avg.Pupil\,Size\,Y \times \pi}{4} \tag{1}$$

$$PCPS = \frac{PS\,Measured\ \text{-}\ PS\,Baseline}{PS\,Baseline} \tag{2}$$

$$APCPS = \frac{\int_{t_1}^{t_2} PCPS(t)dt}{t_2 - t_1} \tag{3}$$

where average pupil sizes X and Y refer to the diameters of a pupil on X and Y axis, respectively, PS measured in Formula (2) is the result of Formula (1), PS baseline refers to the pupil size when a participant relaxes his eyes before the experiment, and t_1 and t_2 refer to the beginning and end of the time window, respectively.

Number of Fixations per Route: Goldberg & Kotval (1998) demonstrated that the number of fixations was correlated negatively to search efficiency. Among the eight 8 routes in the EOP, four were nine-step routes while the rest were shorter routes. Since only successful long routes were analyzed later, the participants might accomplish different numbers of the long routes successfully at the two interface levels. Therefore, the number of fixations per route was calculated for the two interfaces.

Fixation Duration Mean: This indicator was chosen because the fixation duration mean was reported to be correlated positively to difficulty of information abstraction in Goldberg & Kotval (1998).

2.3 Participants

Web advertisements were used to recruit participants. There were 18 male students aged from 19 to 28 (mean= 21.9, SD = 5.8) studying engineering programs at Tsinghua University were recruited as participants.

The participants were required to have normal vision or to wear no more than 3 diopters glasses. Contact lenses were not allowed, due to the use of an eye tracker. They were informed about the details of the experiment protocol and voluntarily signed the informed consent form before the experiment proceeded.

2.4 Apparatus

The experiment scenario was showed in Figure 4. The headband-mounted eye tracker in this experiment comprised with eye- and scene camera assembly which was less than 80 grams. When a participant wore the eye tracker, the position (left or right) of eye- and scene camera in the front of the hat was adjusted according to the dominant eye of the participant. The eye tracker was lightweight so that it was low invasive relatively. Besides, a notebook PC specifically connected to the eye tracker was used for data collection and analysis via a piece of dedicated software.

Fig. 4. Illustration of experiment apparatus

2.5 Training and Pilot Study

To obtain stable operation data, the participants were trained for 2~3 hours to be familiar with the system operation. In the formal test session, all participants were asked to execute the same emergency operating procedure using the two interfaces separately. The eye tracker and its accompanied software collected data of pupil sizes, fixation numbers, fixation durations, blink numbers, and so on.

As a key parameter for data processing, fixation threshold is highly dependent on specific tasks. Therefore, a pilot study was conducted among 5 participants to separate fixation from scanning. The fixation threshold was set at 200 ms consequently.

2.6 Data Analysis

Among the 18 participants, 2 participants adjusted the head-mounted eye tracker arbitrarily and so the calibration was not accurate, thus their data was eliminated. The data of successful long routes of the rest 16 participants were selected for further analysis. All the data of blink number per route, fixation number per route, percentage change of pupil size passed the test of normality. Data for blink rate per minute and fixation duration mean did not pass the normality test. However, for two interface levels, their difference passed the test of normality. Therefore paired t-test with a significance level of 0.05 was adopted for analysis.

3 Results

The results of paired-samples test of the five variables are given in Table 1.

Table 1. Paired Samples Test Results

Variables	Mean	SD	t	Sig. (2-tailed)
Blink number per route	3.2	3.94	3.265	0.005
Blink rate per min	3.26	4.333	3.012	0.009
Fixation number per route	1.8	11.09	0.650	0.525
Fixation duration mean	0.0005	0.0379	0.049	0.962
Percentage change of pupil size	-0.004	0.067	-0.228	0.823

It can be seen from Table 1 that significant difference existed in blink number per route (t=3.265, p=0.005) and blink rate per minute (t=3.012, p=0.009) between the two interfaces. However, no significant difference was found in fixation number per route, fixation duration mean and percentage change of pupil size.

4 Discussion

When the participants performed the experimental task, the blink rate when using the functional interface was significantly lower than that when using the process interface. It indicates that functional interface based on FBTA led to lower workload significantly than the process interface. For the variable APCPS which was supposed to measure workload as well, however, there was no significant difference between the two interfaces. Wang (2011) points out that APCPS may interact with fatigue, nervousness, and other factors, and suggests that APCPS should not be used exclusively.

Number of fixations per route and fixations duration mean showed no significant difference between the two interfaces. This may reveal that the functional interface could not provide higher search efficiency and less difficulty of information abstraction than the process interface. Such non-significance can be explained by several reasons.

First, the experiment task may limit the superiority of functional interface in search efficiency. Components with similar functions were organized together in the functional interface, while similar components may be distributed in the process interface. If a task require a participant to locate a component, the participant can easily locate the region of the component, but then has to carefully check it out of other components with similar functions in the functional interface, while they may spend more time in locating the rough position of the component, but then need to spend little time in finding its exact position in the process interface. Therefore, the difference of the overall search efficiency when using the two interfaces may be not significant. Second, the performance difference between the two interfaces in terms of search efficiency and difficulty of information abstraction may be lessened after the participants were adequately trained. In this study, the participants were trained for 2~3 hours. The significance of performance difference may become too small to be detected with the limited sample size.

5 Conclusion

The influence of interfaces with different information organization on human performance in a simulated EOP task was studied. The findings from this study revealed the importance of designing an appropriate interface for EOP and similar tasks.

The experiment results show that the functional interface based on FBTA was superior significantly to the process interface in user workload. This advantage is of importance. In main control room of nuclear power plants, especially when executing emergency operations, operators are subjected to high pressure. The functional design can reduce workload and may thus further reduce human errors. The results did not prove the superiority of the functional interface to the process one in search efficiency and difficulty of information abstraction, since no significant differences were found in the number of fixations per route and fixations duration mean values.

There were some limitations in this study. (1) that the use of student participants may limit the generalization of the conclusions from this study. (2) The sample size was relatively small.

Acknowledgement. This study was supported by the National Natural Science Foundation of China (Project No. 70931003).

References

1. Adihikari, S., Bayley, C., Bedford, T., Busby, J., Cliffe, A., Geeta, D., Eid, M., French, S., Keshvala, R., Pollard, S., Soane, E., Tracy, D., Wu, S.: Human Reliability Analysis: A review & critique. Manchester Business School Working Paper No. 589 (2009)
2. Brookings, J.B., Wilson, G.F., Swain, C.R.: Psychophysiological responses to changes in workload during simulated air traffic control. Biological Psychology 42, 361–377 (1996)
3. Goldberg, J.H., Kotval, X.P.: Eye movement-based evaluation of the computer interface. In: Kumar, S.K. (ed.) Advances in Occupational Ergonomics and Safety, pp. 529–532. ISO Press, Amsterdam (1998)
4. Iqbal, S.T., Zheng, X.S., Bailey, B.P.: Task-evoked pupillary response to mental workload in human-computer interaction. In: Proceeding of CHI 2004 Extended Abstracts on Human Factors in Computing Systems, pp. 1477–1480 (2004)
5. Kim, J.H., Seong, P.H.: Human Factors Engineering in Large-scale Digital Control Systems. In: Seong, P.H. (ed.) Reliability and Risk Issues in Large Scale Safety-critical Digital Control Systems, p. 163. Springer, London (2009)
6. Rasmussen, J.: The role of hierachical knowledge representation in decision making and system management. IEEE Transactions on System, Man, and Cybermetics SMC-15, 234–243 (1985)
7. Vincent, K.J.: Ecological Interface Design: Progress and Challenges. Human Factors 44, 62–78 (2002)
8. Wu, X.J., Gao, Q., Song, F., Dong, X.L., Li, Z.Z.: Display Design for AP1000 Nuclear Power Plant: FBTA Method vs. Traditional Method. In: The 11th International Conference on Probabilistic Safety Assessment and Management & Annual European Safety and Reliability Conference 2012 (PSAM 2011 & ESREL 2012), Helsinki, Finland, June 25-29 (2012)
9. Wang, Y.: Real-time empirical mental workload measurement in digitalized main control room of nuclear power plant. Bachelors' thesis. Tsinghua University, Beijing (2011)