The 22nd IAMUC

Batumi, 21-22 Oct 2022

Proceedings of the
International Association of Maritime Universities Conference

Head-worn Display Utilization in Engine Supervisory Work

Adi Mas Nizar [1,*], Takashi Miwa [1] and Makoto Uchida [1]

[1] *Graduate School of Maritime Sciences, Kobe University, Japan*
[*] *Corresponding author: 185w604w@stu.kobe-u.ac.jp*

Abstract: Smaller formation of onboard teams demands more efficiency in the task allocation, especially for the engine supervisory work where the operators have to undertake monitoring and maintenance jobs. Besides having sufficient non-technical skills, task allocation can be supported by utilizing cognition aids. We examined the utilization of the head-worn display by measuring the human performance factors such as workload, situation awareness, and trust in automation. A human-subject experiment was conducted using a *2 x 2* within-subject experimental design under a multitasking environment in the engine room simulator. The participants were invited to follow two scenarios: either information on the head-worn display is available or absent. The workload measurement shows that the participants perceived a lower workload when information was available on the head-worn display. It also made the participants put more trust in the alarm system and allowed them to finish more maintenance tasks. However, utilizing the head-worn display bring demerits caused by visual complexity and longer response time to the alarm. We see this trade-off is acceptable to achieve suitable response action. Within the limitation of laboratory-scale experiments, human performance factors promise a valuable result as evaluation method before the head-worn display as a cognition aid is implemented onboard.

Keywords: Situation awareness; trust in automation; workload

1. Introduction

Supporting the onboard operators during the work became more important since team formation is smaller, while the remaining workload is similar and even accumulated to the remaining operators. This issue is also raised by the fact that 80% of accidents in maritime operations are caused by human error (Wróbel 2021); and more specifically, 71% include loss of situation awareness (SA) as one of the causal factors (Grech et al. 2002). Unsuitable work environments also have a role. For example, the recently installed alarm system onboard is described as a nuisance and increases the operators' workload because the system is not appropriately installed (Jones et al. 2006). Therefore, in supporting the recent and future onboard operation, one must include the workload, SA, and trust in the automation as they are an essential aspects to construct the human performance factors (Man et al. 2018).

The working situation of the engine department is generally divided between the engine room where the watchkeeping and maintenance are conducted, and the engine control room where monitoring and administration tasks are undertaken. The scheduled maintenance and watchkeeping leave the engine control room mostly unmanned (M0). When the alarms are activated, the operators need to return to the engine control room to acknowledge and undertake the necessary steps to recover the process back to normal. The issue with the alarm system recently, as its number increased since many sensors were installed, is the increased number of false alarms (Maglić and Zec 2020). The situation makes the operators inefficient in allocating the tasks. To counter this issue, the current engine room situation is supported by an indicator column. It functions as the extended alarm display for the operators to quickly know the sources and categories of the alarms. However, because of the limitation of spaces, it only contained few information than the actual alarm display installed in the engine control room.

The study of head-worn display utilization as a cognition aid is getting attention in various working areas (Bal et al. 2021). For instance, its implementation for patient monitoring promises to increase the SA and reduce the workload (Pascale et al. 2019). In general process control, the same objective can be achieved using a handheld mobile device such as personal digital assistant (PDA). However, it seems less practical in the maintenance tasks since the operators often need both hands to undertake the tasks. Thus, the head-worn display seems more applicable as a cognition aid.

Having information on the head-worn display increases the selection attention because the operators can focus on more essential stimuli (McLaughlin and Byrne 2020). For instance, the alarm display on the head-worn display gives the advantage of providing the raw information continuously for the operator to confirm; it reduces the workload because the operators can discriminate the false alarms and prioritize the tasks (Pascale et al. 2019). However, its utilization is not without disadvantages. The information from the environment might be missed because the display draws too much visual attention. Therefore, this study examined its utilization in supporting operators and evaluated it using human performance factors such as SA, trust in automation, and workload.

2. Methods

Twelve cadet students from the marine engineering department with an average age of 21.7 (±1.1) were invited to participate in this experiment. They have a range of one month to one year of onboard cadet experience. Based on this background, we assumed the participants had enough capability and understanding to undertake the designated tasks in the experiment. The recruitment and experimental procedure were under the code of ethics approved by the faculty board.

We employed a *2 x 2* within-subject experimental design. The first independent variable is the information condition on the head-worn display: the information is displayed (information-on), and there is no information displayed (information-off). The second independent variable is the task-load level differentiated by the number of alarms that annunciated during the scenario: high-task load with twelve alarms and low-task load with six alarms. To counter the limitation of the data acquisition, all participants experienced all four combinations of scenarios.

The experiment was conducted at the engine plant simulator with a separated engine room and the engine control room. The engine room, shown in Figure 1, consists of three monitors to display the mimic diagram of the engine plant and machinery from an actual model ship. The participants can operate machinery and piping systems similar to the actual activity onboard in the engine room, such as opening and closing valves and running the pumps locally. The engine control room consists of the engine control console, as shown in Figure 2, where the participants can monitor the alarm and observe all parameters of the engine and machinery.

We use the Epson Moverio BT-200 shown in Figure 3 as the head-worn display in this experiment. During both information-on and information-off scenarios, the participants can observe the panel displays, as shown in Figure 4, on the engine control console located in the engine control room. However, the participants only have access to the panel display on the head-worn display when the information-on scenario. The display panel consists of six engine system parameters selected using Subject Matter Expert based on the parameter's importance during the engine supervisory work. Each system parameter has three indicators: online process value, high-alarm threshold, and low-alarm threshold.

During the experiment, we introduce two types of alarms: true alarm that activated when the process value increases (or decreases) based on the real engine or machinery trouble, and false alarm that activated because of sensor failure. The false alarm does not need to be handled by the operators. The ratio between true and false

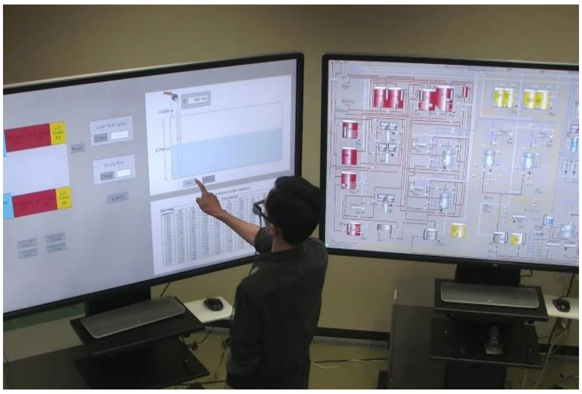

Figure 1 Engine room simulator

Figure 2 Engine control room simulator

Figure 3. Head-worn display device

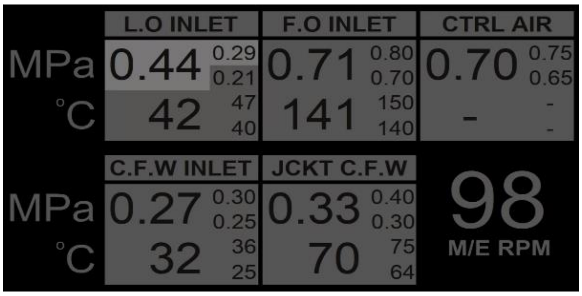

Figure 4. Display image on the engine control console and head-worn display

alarms in one trial was controlled. There were 12 alarms (4 true alarms, 8 false alarms) in the high-load scenario, and 6 alarms (2 true alarms, 4 false alarms) in the low-load scenario. When a true alarm is activated, the process value will slowly increase (or decrease). In contrast, if a false alarm is activated, the process value will suddenly drop (or climb) to the lowest (or highest) level. All activated alarms were followed by indicator color on the display and audible alarms. Whether the participants acknowledge it, the alarm will go off, and the process value automatically returns to normal after 20 seconds.

In order to replicate the work onboard a ship, the participants need to follow two separate tasks simultaneously. The first task came from the engine control room; the participants had several maintenance tasks such as opening, closing the valve, starting or stopping the pump, and sounding several tanks. Simultaneously, participants needed to conduct a monitoring task by acknowledging the alarms from the engine control room when it was activated. Thus, the participants should prioritize the task: decide when to continue the maintenance tasks in the engine room and when to return to the engine control room to handle the alarms.

Before participating in the measurement session, the participants were asked to join the training session three times to get familiarized with the experiment setup layout, the panel on the head-worn display, and the maintenance tasks on the engine room simulator. The training session was followed by the measurement session, consisting of a two-day visit to the simulator. Each day consisted of two trials with different task-load scenarios but the same information status on the head-worn displays. For instance, a participant will have information-on display status on the first day with a high-task load in the first trial, then a low-task load in the second trial. In the second-day experiment, the same participant will follow some procedures with the information-off display status, begin with the low-task load scenario in the first trial, and proceed with the high-task load scenario.

Using a subjective questionnaire, we measured several human performance assessments for the dependent variables: workload, SA, and trust in automation. We used NASA-TLX (NASA Task Load Index) to measure perceived workload (Hart and Staveland 1988). It consisted of 20 Likert-type scales constructed by six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. The perceived workload was then calculated by summing up the six dimensions after weighing it with factors from the participants. SA is measured with subjective SART (Situation Awareness Rating Technique) (Taylor 1990). It covers ten dimensions grouped into three categories: attentional demand, attentional supply, and situational understanding. The perceived SA is simply defined as the difference between attentional demand and attentional supply that eliminated from the total understanding of the situation. The trust in the alarm system is measured using the TiA (Trust in Automation) questionnaire with 12 dimensions (Jian et al. 2000).

The objective measurement was done by comparing the number of maintenance tasks conducted by the participants in one trial. We also recorded the response time (*RT*) in seconds correction for all activated alarms during the one trial. When participants neglected the alarm, the response time was the maximum since it is automatically deactivated at $RT = 20$ seconds. The measurement also compares the ratio of participants' actions to acknowledge or neglect the alarm.

3. Result

The perceived workload measured using NASA-TLX was tested with two-way ANOVA to see the interference of display status and alarm frequency. The representative of the result is shown in Figure 5. As the result there was no two-way interferences ($F(1,11) = 1.63$, $p = .23$). However, the analysis for each variable bring the result of significance ($F(1,11) = 7.96$, $p < .05$) on the display status effect: participants perceived more workload during the information-off trials ($M = 11.43$, $SD = 2.57$) rather than in the information-on trials (M =

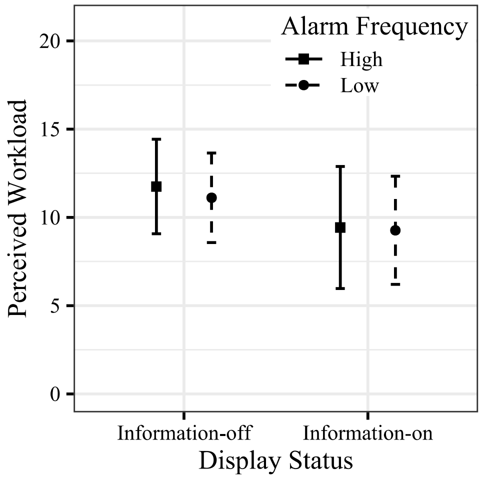

Figure 5 Interference of display status and alarm frequency on perceived workload

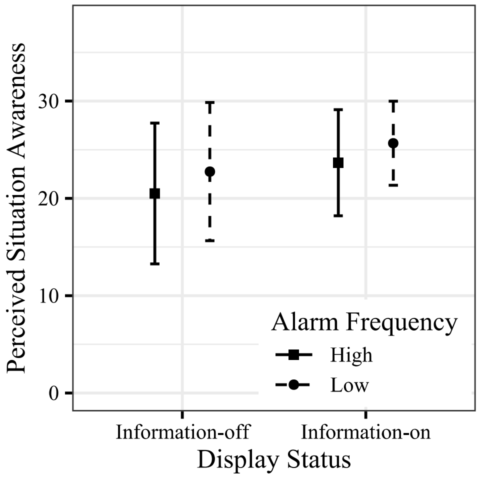

Figure 6 Interference of display status and alarm frequency on perceived situation awareness

9.35, *SD* = 3.20). However, the alarm frequency interference did not give any different for both conditions (*F*(1,11) = 1.34, *p* = .27).

The perceived SA that measured using SART given the result shown in the Figure 6. There was no two-way interaction developed (*F*(1,11) = 0.04, *p* = .85), tested with two-way ANOVA. The display status effect also does not give a significant difference (*F*(1,11) = 2.58, *p* = .14). However, with different alarm frequency, the participants show higher SA in the low alarm frequency session (*M* = 24.21, *SD* = 5.94) than in the high alarm frequency session (*M* = 22.08, *SD* = 6.47), tested with statistically significant difference (*F*(1,11) = 6.97, *p* < .05).

The subjective questionnaire result from TiA that measured trust in alarm system indicate participants perceived more trust on the alarm system in the trials with the information-on (*M* = 31.21, *SD* = 7.01) than in trials with the information-off (*M* = 27.75, *SD* = 8.39), with statistically significant difference (*F*(1,11) = 7.32 , *p* < .05). Meanwhile there was no interference effect of alarm frequency (*F*(1,11) = 0.33, *p* = .58), nor two-way interference include the display status effect (*F*(1,11) = 0.55, *p* = .47) that tested with two-way ANOVA. The interaction between variables on the result is shown in Figure 7.

The objective measurement result of total completed tasks in one trial is shown in Figure 8. Two-way ANOVA was conducted to analyze the result. There was no two-way interaction between the variables (*F*(1,11) = 0.38, *p* = .55). However, the participants finished more task in the information-on trial (*M* = 122.13, *SD* = 27.61) compare when they were under information-off trial (*M* = 109.08, *SD* = 23.11), but with less statistically significant difference (*F*(1,11) = 12.97, *p* = .07).

The objective measurement regarding the alarm response is shown in Figure 9 for the alarm response ratio. There was two-way interaction between the variables of display status and alarm category (*F*(1,11) = 41.18, *p* < .01). The *post hoc* t-test for each variable shows that the participants responded less often to the false alarm

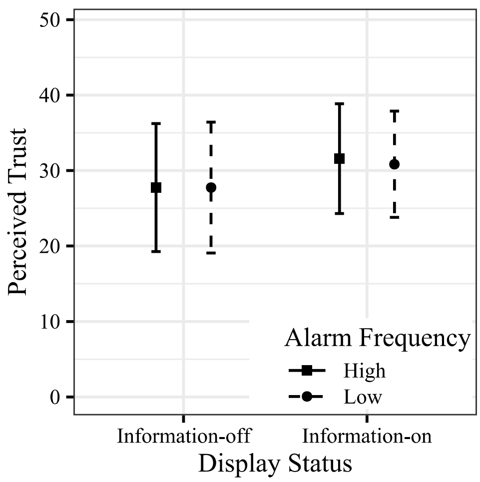

Figure 7 Interference of display status and alarm frequency on perceived trust

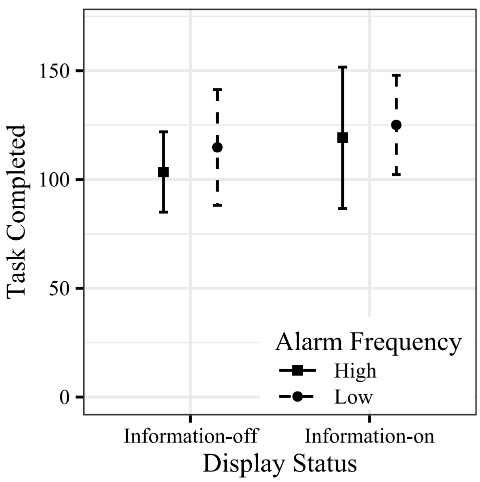

Figure 8 Interference of display status and alarm frequency on the task completed

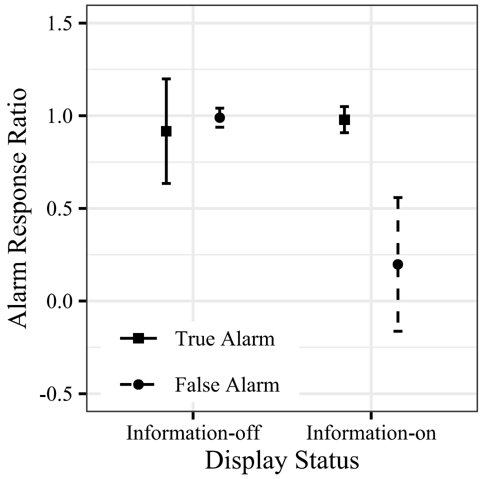

Figure 9 Interference of display status and alarm category on alarm response ration

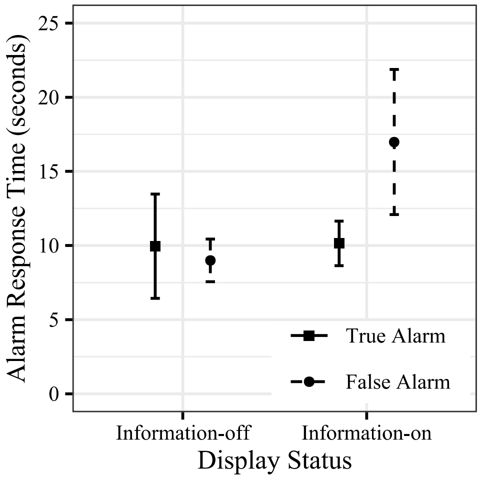

Figure 10 Interference of display status and alarm category on alarm response time

during the trial with the condition of display status information-on (M = 0.97, SD = 0.36) compared to when they were in the trial with information-off status (M = 0.98, SD = 0.51); tested with a statistically significant difference (p < .01).

Another objective measurement came from alarm response time, measured in seconds correction, as shown in Figure 10. There was no two-way interaction between variables, tested with two-way ANOVA (F(1,11) = 0.51, p = .49). Also, the interference on the single effect of the alarm frequency did not give the significance effect (F(1,11) = 0.99, p = .34). However, the participant responded to the alarm slower during the trial with the information-on display status (M = 10.27, SD = 1.48) compare when they were conducting the trials with the information-off display status (M = 8.94, SD = 1.27), tested with statistically significant difference (F(1,11) = 7.69, p < .05).

4. Discussion

An experiment conducted in this paper examined the head-worn display utilization in the engine supervisory control onboard a ship. Using the engine plant simulator, cadets were invited as the participants. During the experiment, they were introduced to the two head-worn display setups (information-on and information-off) and two task-load conditions differentiated by the alarm frequencies (high-task load and low-task load). The objective measurements such as the number of finished tasks and response time to the alarm are recorded beside the subjective measurements such as workload, SA, and trust in automation questionnaires.

The experiment findings indicate that the participants perceived a lower workload in the information-on trials. It suggests that when the participants have information to confirm an incoming alarm, whether true or false, they can safely neglect to return to ECR. Furthermore, the participants perceived a lower workload by moving between the engine room and engine control room less frequently. However, in this experiment, the alarm frequency did not affect different perceived workloads for the participants. Therefore, it makes us unable to examine the correlation between workload and SA with trust in automation.

The SA, one of the human performance factors that we predicted would increase if participants had information on the head-worn display, did not make any difference. It explains as a trade-off between attention demand and attention supply. While attention supply does increase with the information on the head-worn display, attention demand also increases because the layer of information makes the visual environment more complicated. Therefore, continuously putting information on the visual display is less effective than we thought because it did not improve the SA. Based on this consideration, in practice, information on the head-worn display should be switched off when the engine parameters are in average running condition and automatically switched on when the parameters are developed towards the alarm threshold range.

Perceived trust in the alarm system measured using subjective measurement explains that the participants put more trust in the alarm system when they wear a head-worn display with additional information. Moreover, in the same condition, the response ratio comparison result shows that the participants respond to ECR less often if the false are were activated. Contrary, participants might adopt a "better-safe-than-sorry" approach in the trade-off between maintenance and monitoring tasks when no raw information is available; they took the safe course of action and returned to the ECR every time the alarm sounded.

There is an indication of task allocation improvement since participants finished more maintenance tasks when information was available on the head-worn display. Furthermore, with information available to confirm the false alarms and participants putting more trust in the alarm system, they could safely neglect the false alarms from the ECR and make prioritization between multitasks efficiently. Meanwhile, the response time showed a different tendency: participants with information on their head-worn displays took longer to respond to the alarms. The participants said they only confirmed the information after the alarm sounded, and this confirmation length made the response time longer compare to when there was no information to confirm. Nonetheless, a suitable response action in the engine supervisory control is more important than a quick response time. We suggest that this trade-off between confirmation and response time is acceptable for better task allocation.

There is a limitation in the experimental setup where the information on the head-worn display was continuously present. One must consider including the setup where the participants could activate or deactivate the information presented on the head-worn display. The limitation also comes from the experimental setup where the study was conducted in the simulator using screen displays. It is advisable to conduct the experiment at the environment without screen display to eliminate visual environment complexity when using a head-worn display. However, with the limitation of the laboratory scale experiment, several human performance factors such as workload, SA, and trust in automation show a promising result as an evaluation method before the new cognition aid, such as the head-worn display in this study, is implemented onboard.

Acknowledgments

This research was funded by JSPS KAKENHI Grant No. 21K04517. We want to thank Dr. Tatsuro Ishida for the feedback during scenario construction and Mr. Daiichi Aburagi for technical assistance at the engine plant simulator.

References

[1] Bal M, Benders J, Dhondt S, Vermeerbergen L (2021) Head-worn displays and job content: A systematic literature review. Appl Ergon 91:103285. https://doi.org/10.1016/j.apergo.2020.103285

[2] Grech MR, Horberry T, Smith A (2002) Human Error in Maritime Operations: Analyses of Accident Reports Using the Leximancer Tool. Proc Hum Factors Ergon Soc Annu Meet 46(19):1718–1721. https://doi.org/10.1177/154193120204601906

[3] Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In: Advances in Psychology. Elsevier, https://doi.org/10.1016/S0166-4115(08)62386-9, pp 139–183

[4] Jian J-Y, Bisantz AM, Drury CG (2000) Foundations for an Empirically Determined Scale of Trust in Automated Systems. Int J Cogn Ergon 4(1):53–71. https://doi.org/10.1207/S15327566IJCE0401_04

[5] Jones BS, Earthy JV, Fort Ed, Gould D (2006) Improving the design and management of alarm systems. In: World Maritime Technology Conference. IMarEST, 2006, p 7

[6] Maglić L, Zec D (2020) The Impact of Bridge Alerts on Navigating Officers. J Navig 73(2):421–432. https://doi.org/10.1017/S0373463319000687

[7] Man Y, Lundh M, MacKinnon SN (2018) Managing unruly technologies in the engine control room: from problem patching to an architectural thinking and standardization. WMU J Marit Aff 17(4):497–519. https://doi.org/10.1007/s13437-018-0159-y

[8] McLaughlin AC, Byrne VE (2020) A Fundamental Cognitive Taxonomy for Cognition Aids. Hum Factors J Hum Factors Ergon Soc 62(6):865–873. https://doi.org/10.1177/0018720820920099

[9] Pascale MT, Sanderson P, Liu D, Mohamed I, Brecknell B, Loeb RG (2019) The Impact of Head-Worn Displays on Strategic Alarm Management and Situation Awareness. Hum Factors J Hum Factors Ergon Soc 61(4):537–563. https://doi.org/10.1177/0018720818814969

[10] Taylor RM (1990) Situational Awareness Rating Technique (SART): The Development of A Tool for Aircrew Systems Design. In: Situational awareness in aerospace opeartions (AGARD-CP-478). NATO-AGARD, Neuilly Sur Seine, France, p 3/1-3/17

[11] Wróbel K (2021) Searching for the origins of the myth: 80% human error impact on maritime safety. Reliab Eng Syst Saf 216:107942. https://doi.org/10.1016/j.ress.2021.107942