Situation Awareness in Engine Resources Management :Its interplay in cognition, reliability, and socio-technical system

Nizar, Adi Mas

(Degree)
博士（工学）
(Date of Degree)
2023-09-25
(Date of Publication)
2025-09-25
(Resource Type)
doctoral thesis
(Report Number)
甲第8747号
(URL)
https://hdl.handle.net/20.500.14094/0100485931

Doctoral Dissertation

Situation Awareness in Engine Resources Management : Its interplay in cognition, reliability, and socio-technical system

（船舶機関管理における状況認識に関する研究：
認知、信頼性、社会技術システムとの相関）

July 2023

Graduate School of Maritime Sciences
Kobe University

ADI MAS NIZAR

Contents

List of Figures

List of Tables

Publications

Journal

- Nizar, A.M., Miwa, T., Ozaki, T., Uchida, M. *Aiming for Alarm System Improvements for Ship Engine Operation Using Engine Control Console Alarm Data. Marine Engineering*. Vol. 56 No.6 p.961-969. 2021.
- Nizar, A.M., Miwa, T., Uchida, M. *Measurement of Situation Awareness in Engine Control Room: Approach for Non-technical Skill Assessment in Engine Resource Management*. WMU Journal of Maritime Affairs. Vol. 21 No. 3 p.401-419. 2022.
- Nizar, A.M., Miwa, T., Uchida, M. *Examining Situational Awareness, Trust in Automation, and Workload in Engine Resources Management: An Evaluation of Head-Worn Display Technology*. WMU Journal of Maritime Affairs. 2023.

Conference

- Nizar, A.M., Miwa, T., Uchida, M. *Measuring Situation Awareness in Engine Control Operation*. International Association of Maritime Universities (IAMU) Conference 2021. Alexandria, Egypt (online).
- Nizar, A.M., Nakamura, M., Miwa, T., Uchida, M. *Dynamic Safety Modelling for Ship Management Performance*. Maritime Safety International Conference 2022. Surabaya, Indonesia (online).
- Nizar, A.M., Miwa, T., Uchida, M. *Head-worn Display Utilization in Engine Supervisory Work*. International Association of Maritime Universities (IAMU) Conference 2022. Batumi, Georgia.
- Nizar, A.M., Miwa, T., Uchida, M. *Human-Machine Interface Evaluation in Engine Supervisory Control through Alarm Performance Assessment*. (IMarEST) International Ship Control System Symposium 2022. Delft, Netherlands.

Chapter 1

Background

1.1 Human element in maritime operation

The contribution of humans in maritime accidents, approximately number is around 80%, has been consistently highlighted that human involvement in the accident or incident plays a significant role. The same ratio is often found in other industries besides the maritime operation [2]. While this number may be subject to debate in the recent decade, a comprehensive review study conducted by Wrobel concludes that human error remains a substantial factor in maritime accidents [3]. The recent circulation keep the importance of addressing human factors to improve safety measures within the maritime industry, not only ended to blame the operator as the source of accident.

Recognizing the influence of human actions, the International Maritime Organization (IMO) use the term *human element* to explain aspects related to human involvement in the maritime industry, not only during the operation, but also within its design and evaluation. The factors that are noted in the human elements include knowledge, skill, experience, decision-making, situational awareness, communication, and behavior. By acknowledging the complex interplay between human actions, organizational structure and technical system, a comprehensive approach to maritime safety that includes human element or human factors is necessary with robust and heuristic.

One that listed in the human element by IMO is situational awareness, or interchangeable term with situation awareness in general field of human factors. Looking back to accident reports, situation awareness in the maritime operation is amplified since it contributes to the 71% of human error [4]. Situation awareness in general and at the human individual level, refers to the ability of individuals to perceive, comprehend, and anticipate the dynamic elements of their environment in relation to their goals and objectives. The importance of situation awareness is amplified as it is included in the non-technical skill in the crew resources management [5].

For most of the studies in the human factor or human error, the most focus is given for navigational work or bridge department, compared to the engine operation. This issue can be understood since the concept of safety and human factors in maritime operation, including its crew resources management, is adopted from the studies and best practice in the aviation industry [6]. However, it is important to acknowledge the significance of the human factors in the engine department as well. Engine operations entail their own set of challenges and potential sources of human error. Engine operation, including the engine supervisory control, is generalized as a process control where the vigilance task with long duration and its monitoring task of process parameters reading [7, 8]. Looking at the situation of studies in the maritime operation, more studies are needed to balance the studies on the navigational work to ensure holistic understanding of the human factors or human elements also include the environment work of engine operation.

1.2 Human-machine interface issue

The recent issue regarding human factors in engine operation is with its human-machine interface (HMI), especially in engine control console side. As automation and technological advancements continue to progress, the task that performed in the engine control room (ECR) have become more demanding due to the utilization of recent information source for process control [9]. While automation reduces the active involvement of engineers into the system, it still allows the human to remain in the main control circuit [10]. However, the combined pressure of reduced crew members numbers work onboard and increased automation have significantly transformed the role of operators. It makes a trade-off between automation and operator effectiveness [4]. When automation is introduced, the operator's monitoring capabilities are compromised [11]. There is a drawback when the condition is in the normal operation as the operator workload will increase because the operator needs more time and attention to gather the situation awareness of recent operation carried out by the automation.

With the progression of advanced automation, the role of HMI that facilitates the warning or alarm system is becoming crucial. The reported problems with the installed alarm system recently, for instance, the installed alarm system is considered to be overly prioritized during the monitoring [12]. Some engineers on board, in another study also noted that the recent alarm system, in the form of list, are subject to complaint [13]. Moreover, the concern also includes nuisance alarms that distract operators from their current task. The lack of alarm messages clarifies is also confusing the operators and makes the excessive workload.

The alarm data that comes simple in dataset with alarm name, type, id, activated time, and deactivate time is the source of alarm performance assessment in the scope of alarm management [14]. Similar guidelines in maritime operation can be found within the IMO Code on Alerts and Indicators 2009 [15]. The data from a three-year voyage of training ship and three months voyage of ocean-going ship is revealed the recent issue with the alarms system onboard a ship. The two most mentioned in the recent studies in alarm management, also appear in these data. They are false alarm and alarm floods problem.

Alarm flood is defined as the condition where multiple alarms are activated at the same time during short period [16]. During the state of alarm flood, the engineer on board may find difficulty in handling the alarm. Since the guideline in the maritime operation does not regulate nor give steps to quantify the alarm floods, the guide onshore processing plant is used as a reference. For instance, alarm floods here defined as the condition where ten alarms or more are activated during the ten minutes. A good example is found in Figure 1.1 where there are 13 alarms floods from the ocean-going vessel, where the horizontal axis explains 10 minutes bin of the operation and indicated how many alarm is activated during its bin. The recent prominent alarm management is needed to bring the number of this alarm to the lowest point. Taking the example of the effect of alarm floods in bad condition is with Emma Maersk's engine room floods [17]. Although the trouble comes from mechanical breakdown, the condition is getting worse because the engine control room at that time has another *flood* from the alarm that keeps coming. This condition makes the engineer on the watch not have enough resources to attend the real water flood at the engine room. These examples also give the statement that alarm floods may be neglected in ship operation, but it becomes a dormant factor when it combines with the worse condition in the engine operation.

The second most problem with the alarm system is false alarms, or in another terms nuisance alarms. It became a nuisance since the response from the operators is not needed. Most of them are chattering alarms where the same alarm keep repeated between activated and deactivated [14]. These alarm types may be activated by only a single error that operates close to the alarm threshold [18]. The Fig 1.2 left part shows list of alarm when it compared with its deactivated time under 10 seconds. It is confirmed with the engineer onboard: the alarm such M/E starting air pressure is not necessary occurred during leaving port since the engineer well understood that the air compressor is just running, as illustrated in Fig 1.2 right part. The appearance of such nuisance alarm,

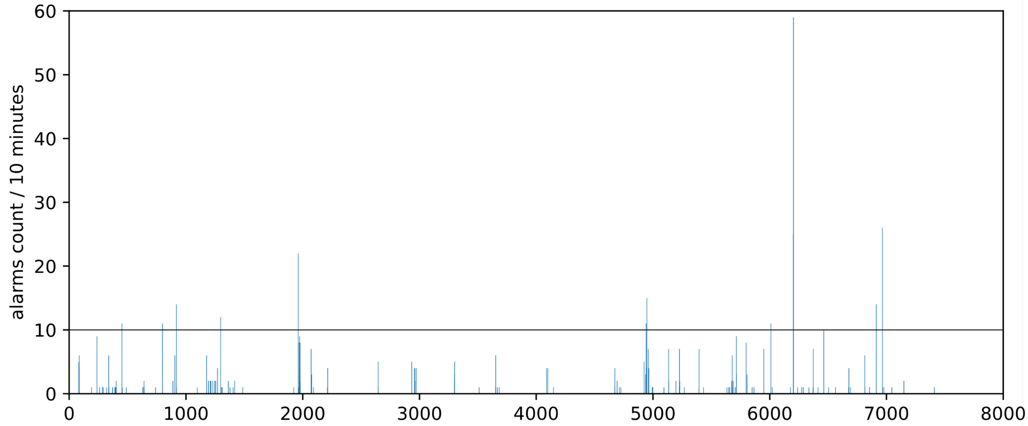
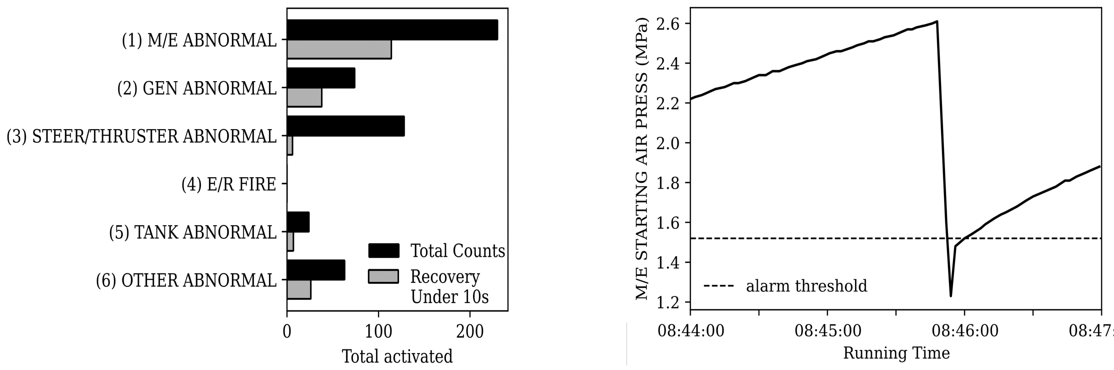


Figure 1.1: Alarm floods from a three months ocean-going ship

furthermore, can reduce the performance of engineers in the first step in handling the malfunction situation. In various workload conditions, the engineer onboard may have reaction time and accuracy may that declined, or in the extended condition is to silence the alarm. Moreover, in the long exposure of nuisance alarms, it stimulates the cry-wolf syndrome where the engineer onboard trust is reduced.



(a) (b)

Figure 1.2: False alarm ratio and sensor reading from a traning ship

Various counter problems in alarm management were introduced based on best practice explained in the guideline. To counter the nuisance alarm for instance, can be done by introducing a delay time before alarm activated in serial activation or alarm set-point is changed to lower/higher from the previous, this practice is introduced with the recent data, and the result has been seen in reducing several nuisance alarms as shown in Fig. 1.2. In case the alarm floods condition, a recommended approach is to offer operators a decision aid, such as alarm prioritization, for instance as the best practice is to prioritized alarm into low, medium, high priority with proportion of 80%, 15%, and 5% respectively. Many studies also approach by enhancement on the alarm display to guide the operator in identifying the source of the problem and determine which alarm requires their immediate attention.

1.3 Aim of Research

The general aim of the study is to examine the concept of situation awareness in engine resources management. Situation awareness that has been explored in maritime operation is less comprehensively defined in the engine operation, compared to the navigational department. The recent issue in HMI becomes the starting point of the research to address situation awareness in the scope of human cognition how the operators construct the situation awareness and what factors that influence it. The concept of situation awareness is also addressed in the scope of human reliability to what extent the situation awareness became the latent factor to the human error and how to measure it. Lastly, the concept of situation awareness is also examined in wider scope of the relation between onboard engineer and onshore support management where the situation awareness plays important roles in maintain well engine resources management.

1.4 Outline of studies

With three scopes namely in human cognition, human reliability, and socio-technical system, this research employes different framework for each study. The outcomes of the study also differ but mainly is recent framework improvement and its sensitivity in specific engine resources management.

Study I Human cognition. The close HMI issue remains situation awareness as an important concept. Within the human cognition scope, situation awareness is measured and examined its interplay with other human performance factors, in this study are workload and trust in automation. The engine plant simulator is used as the medium for conducting experiments with human factors framework. The various measurements from subjective using questionnaire, objective using the freeze-probe and measurement point from engine plant simulator are employed in this study. Situation awareness is also examined in the case of additional cognition aid being introduced to the work area.

Study II Human reliability. Human reliability analysis is often constructed solely from the accidents or incidents report. Within this study, the approach using experimental results as the bottom-up methods is applied to predict human error probability. The performance shaping factors that previously treated dependently are calculated using Bayesian network. To reduce uncertainty, the data from simulator experiment is used alongside the input from the expert judgment. Different stages of node are constructed based on its relevancy in the simulator experiment data.

Study III Socio-technical system. Another issue from alarm performance assessment is this practice has not existed in recent maritime operation. Thus, lack of understanding between onboard engineer and onboard management may exist. This study aims to examine the situation awareness interplay with safety, efficiency, and workload. To model the socio-technical system between onboard and shore management, dynamic safety theory simulated using System dynamics. The observation is made for the operation behaves following the safety and efficiency pressure, that created by existing goals and boundaries.

Chapter 2

Situation awareness in human cognition

2.1 Information processing model

In the cognition level, situation awareness has been studied extensively in various fields, including in maritime operation, to understand how individuals perceive and understand their work environment. Situation awareness has been defined in many forms, regarding the level and the scope. It depends in what environment the human should conduct the work. In the cognitive level, the most adapted definition could be came from the Endsley who defined the situation awareness as *"the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and a projection of their status in the near future"* [19]. Based on this definition, there are three levels of situation awareness started with perception stage where the operator perceived information that important to the task, comprehension stage where the integration of information done to understand the meaning of recent condition, and the projection stage where the operator can predict the condition will be going in the future.

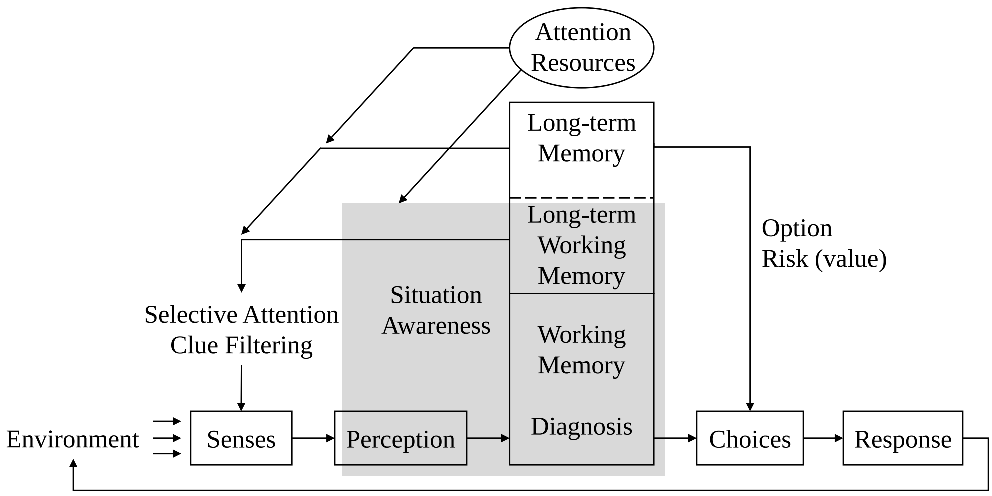


Figure 2.1: Information processing flow, from [1]

The human cognitive study is based on the human information processing stage [1], as shown in Fig 2.1. This model is similar to the model in process control, where the main stage is input, process, and output as the feedback. The process starts at the sensory processing where the cues are consumed from the environment. These include utilization of five human senses, where most of the human information processing comes from the visual. In the context of engine operation, this cue came from alarm sound, engine parameters at the consoles, engine vibration, and radio communication. The next step is perception where these cues are interpreted to the brain. It gives a meaning to the information so the human can convey such meaningful information. The role of selective attention then to integrate this information from working memory with the long-term working memory. Such process includes to choose which cues to be processed to understand recent and future state of the engine operation. The hypothesis triggers several action options in the decision making stage. The feedback from the action will influence the system where the human interacts with.

Situation awareness, as explained in the information processing flow above, has a role in combination process of perception, working memory and long-term memory [1]. This made by the function of attention resource that prioritize the clue that will be include in the perpception stage. In the cognition level, the situation awareness often made by the incorrect cues, the lack in hypothesis construction, a drawback in the selective attention, and working memory. In the ship engine operation also interpret these into: error in information is not registered, information is registered but not processed, and information registered but was illuded [20]. Thus, the function of situation awareness is believed to be important in helping the operator to separate the information to only that they need to execute the specific aim [21]. However, situation awareness is limited and not include decision-making and action which choosing among different options and carrying out a chosen course of action. However, situation awareness roles are focused on maintaining an accurate understanding of the environment, which further can support effective decision-making and action, for sure combined with other factors [22].

This situation awareness by its definition, further, cannot be directly translated that it is sequently linear stages [23]. Instead, situation awareness is a dynamic that involves perception and comprehension of the current situation, leading to a higher situation awareess. For instance, the operator who maintains the comprehension stage has better situation awareness than them who can only maintain the perception stage. As an iterative model, situation awareness in human cognition may drive the operator to actively search for information to feed the comprehension or their perception stage [24]. Situation awareness is constantly evolving and an adaptive process that requires ongoing attention and effort from the operator to maintain.

To process the information, for example in the case of engine supervisory control, it relies on long-endurance work. For such, workload is important factor to influence how situation awareness is constructed [23]. High workload can lead to lowering SA, as the operator ability to process information is became overwhelmed [25]. A higher workload can also cause the operator to misprioritize information, leading to information being overlooked or ignored. The appearance automation to reduce workload also has other issues as the trust in automation. This situation awareness, workload, and trust in automation became the prominent issue in human factors recently [26]. These also became generally used to evaluate human-machine interface in process control environment [27].

2.2 Human factor approach

Human factors is defined as scientific discipline with concern in the interaction between human and other elements on the system to optimize human wellbeing and system performance [28]. As explained before, IMO prefers to use term *human element* to explain the factors that include human in maritime operation [29]. The human factors in maritime became important look back in situation that 80% accidents occurred because the human error [4]. Since the human factors became important especially after some accident occurred, the implementation of regulatory and classification aspect became prominent in design stage and operation. However, much of study falls only in the analysis of accidents [30]. The limited capability of it cannot capture the human factors in the routine work on board a ship. The accident analysis alone will only blame the human as the cause of accident or near misses but neglect their role to prevent it happen in the first. To approach system safety and to examine the cause behind the problem, quasi-experimental is needed to approach everyday routine work [30].

Measurement of SA in human cognition level can be a direct measurement to its process, or infering it from observable processes, behaviors, or performance. The indirect measurement has the most flexible and convenience measurement by putting the questionarie in the of trial. Despite the potential advantages of indirect measures, direct measurement is considered the more effective approach, as it provides a more accurate representation of situation awareness measurement [31]. However, the direct measurement has a disadvantage of being intrusive to the trial that makes the extra scheduling to put the measurement during the trial. The direct measurement in further, can be divided into subject and objective measurement.

For the objective measurement, the global assessment technique (SAGAT) was introduced firstly to apply the freeze-probe technique [19]. The derivative of this freeze-probe in the process control environment or nuclear power plants was introduced as situation awareness control room inventory (SACRI) [32]. Both methods required the participant during the experiment to pause the scenario and needed to answer several questions regarding ongoing operation. The difference is, SAGAT is a more general approach method where the construction of queries is not provided as detail in SACRI. The freeze-probe requires the observer to design and understand the scenario that wants to be evaluated. Thus, this technique requires domain-based understanding about the experiment [33]. In maritime operation, the freeze-probe were evaluated for instance in training of collision avoidance [34]. The result of modified SAGAT shows the ability to discriminate the effect of training to recognized number of vessels. In hardware development, freeze-probe was also evaluated to be sensitive to find the effectiveness of integrated navigation system (INS) compared to the recent bridge [35].

In subjective measurement methods, situation awareness rating technique (SART) is the most notable method [21]. Subjective measurement in general needs the participant to answer the perceived situation awareness [36]. These methods are noted as convenient and easy to use but have a drawback compared to objective measurement. This is related to the participants' beliefs rather than realized in what circumstance their accuracy is. However, its general construct has an advantage to be applied in wide areas of research. Thus also, the subjective measurement should be considered as complete to the objective measurement [37].

Situation awareness is critical for effective decision-making, especially in high stress and complex environments [21]. Two other human performance factors that influenced this came from the workload and trust in automation [26]. High workload can reduce situation awareness by limiting the ability of individuals to process and integrate information. Trust in automation that is not well maintained will lead to situation awareness loss. Thus, it is important to consider workload and trust in automation when designing and implementing situation awareness study. Situation awareness, human workload, and trust in automation then together form the measurement are used to evaluate the human-machine interface, especially in process control [27].

2.3 Cognition aid to improve work performance

A higher level of situation awareness is achieved when the operator can comprehend the situation and anticipate its future development, as opposed to only perceiving data [23]. The improvement of situation awareness, for instance, can be done by provide training and cognition aid. To improve situation awareness, it is necessary to provide operators with the necessary information, context, and training to comprehend the situation accurately and quickly. Conversely, two approaches to furher determine situation awareness as the concept in developing the situation awareness oriented training and situation awareness as te way to evaluate the design concept.

Using a head-worn display can help operators focus on essential stimuli and perform attention-requiring tasks, improving situational awareness and reducing the need for social support. This cognitive aid can store information and provide continuous display, allowing operators to access data without relying on confirmation from other operators in another location. This can enhance the work environment and drastically change the way tasks are performed. However, using a head-worn display may also have drawbacks, such as misfocus, increased workload, and disruption of the learning curve. Operators may also fear that their skills may become obsolete.

In supervisory control, it is crucial to have both data-driven and knowledge-driven monitoring. While head-worn displays can help reduce false alarms and provide contextual information before an alarm sounds, they can also increase the workload when information saturation occurs. The information on the head-worn display can simplify the decision-making process and encourage better alarm prioritization. However, operators must also consider the trade-off between having access to information and the potential disruption of their skill acquisition process. Overall, while head-worn displays have the potential to improve performance and situational awareness, operators must carefully weigh the advantages and disadvantages before using them in their work environment.

2.4 Situation awareness measurement in engine room situation

The first experiment aims to measure the situation awareness in an engine supervisory environment. The literature recently only supplies the examination of the SA in the environment of navigation deck. Thus, within this experiment the sensitivity of SA understanding was considered. Beside the situation awareness alone, its interaction with workload as the main correlation factor in shaping situation awareness was also examined in this study. Another factor that is examined is the participants' background regarding their training experience on board a ship.

Quasi-experimental using full-mission engine plant was done. The experimental design was mixed *2x2x3*. The first factor is made by two-level scenarios: entering port (EP) and ocean-going (OG). The intention is to create two different workload environments. Second, it contains two levels of participation: first time (E1) and second time (E2), since the participants joined the experiment two times. Last, the design contains quasi-experimental independent variable of participants background on the training ship duration: one month (1M), two months (2M), and three months (3M).

Twenty-four participants (age average 21.8 $\pm$1.1) consist of undergraduate and graduate students belonging to the marine engineering course have joined the experiments. The requirement was made based on the cadet training ship duration; it is able to divide evenly the participants into three groups. For conducting the experimental procedure, the participants were assumed to have the same baseline knowledge of using the engine plant simulator. The difference between participants was based solely on the onboard training experience. The experiment was designed and performed under compliance to the code from the faculty board. Before the experiment is conducted, the participants give the informed consent and sign the agreement to join the experiment.

2.4.1 Experimental setup

High-fidelity full-mission engine plant simulator was used to conduct the experiment. In this first experiment, only engine control room (ECR) was used, as illustrated in the Figure 2.2. The engine control console is divided into three sections: the left section contains analog parameters, the middle section contains the engine mimic and digital parameters, and the right section has a group starting panel (GSP). Two scenarios were created using the simulator function to record and replay the created scenario. The idea of using the record-and-replay modes also to prevent the participants action from altering the process value in engine operations. Its ensure that all participants experienced the same exposure and timestamp for malfunction events.

In order to replicate the actual working environment of engine operations, entering port and ocean-going scenario were developed to examine the impact of task-load on situation awareness. The entering port scenario required the participants to follow a stand-by engine procedure and manage a larger task-load. This was created by confirming the parameters changes caused by fictional companion engineer actions and reported by recorded voice, such as manually starting

Figure 2.2: Engine plant simulator layout, ECR side

the compressor and changing fuel supply. Participants also had to adhere to the communications from both engine room and recorded officer at bridge during the confirmation. However, this procedure was not needed in an ocean-going scenario. For both scenarios, there were two different sets of malfunctions, which were cross exchanged between trials on the different days of the participation (E1 and E2). This was done to prevent the participants from becoming familiar with the engine control console parameters included in the freeze-probe questions. During the scenario, the participants were required to acknowledge each alarm caused by a malfunction by pressing the stop button on the engine control console and reporting the alarm messages to fictional companion engineer in the engine room by using radio communication.

Participants were scheduled to attend the simulator three times on the three separated days: briefing days, first time participation (E1), and secondtime participation (E2). During the briefing day, the purpose and setup of the experiment were presented, along with training on the stand-by engine procedure. On the first time participation day, participants underwent two trials of two different scenarios. Lastly, during the second time participation day, the same scenario was repeated with two trials, but in a different sequence.

2.4.2 Measurement instrument

The freeze-probe technique was modified based on the SAGAT to produce measurement that is more direct and objective. However, this technique has the drawback of being disruptive to the task performance. To implement the freeze-probe technique, the simulator was paused several times during the experiment, and the participants were instructed to stop their activity and answer a set of questions related to the ongoing scenario. To prevent participants from being cued by the parameters on the engine control console, they were asked to turn away from it while answering the questions.

The timing of the simulator freezes, and malfunction was predetermined for each scenario as shown in Figure 2.3, but the information was undisclosed to the participants. This prevents the participants monitored the parameters only when approaching a simulator freeze. As the experiment design, there were three simulators freeze per trial, each lasting less than five minutes, to comply with the previous methods /citeendsley2012. A malfunction-sensitive and scenario-sensitive categorization was used to construct the eight questions asked during each simulator freeze, refer to process view technique [38]. In more detail, the malfunction-sensitive questions related to a fault that was activated before the simulator paused, while the scenario-sensitive questions were unrelated to the fault but were relevant to the scenario context. The difference with the SAGAT is the questions were not randomly assigned as in the SAGAT procedure. The questions were assigned based on the top-down sampling approach in the monitoring process.

Within the questions for each freeze, they are categorize into situation awareness perception stage question, and comprehension stage question. Perception questions were constructed based

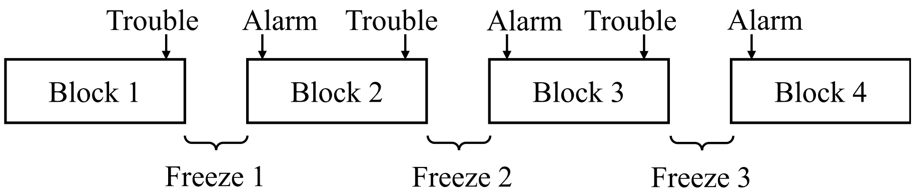


Figure 2.3: Timeline for malfunction and freeze-probe

on the process parameters displayed on the engine control console, while the comprehension questions were not indicated, but the participant can answer based on the displayed parameters. Discrimination of each stage is shown in Table 2.1. A total of twenty-four questions were created for one trial, with an equal proportion of perception level and comprehension level questions, as well as malfunction-sensitive and scenario-sensitive questions. Rather than comparing the current position to the past position, the questions were designed to compare the current position to a reference point in order to avoid confusion. This was done to eliminate any ambiguity in the correct response due to differences in the interpretation of normal process control conditions. For instance, during engine start-up, the compressed air pressure typically decreases, and as such, a pressure drop should be anticipated under those circumstances.

Table 2.1: Freeze-probe question, P for perception, and C for comprehension

	Question
P	Compare with past, how was F.O INLET TEMPERATURE developed?
C	Compare with past, how was F.O INLET VISCOSITY developed?
P	Compare with past, how was CENTRAL C.F.W PUMP INLET TEMPERATURE developed?
C	Compare with past, how was M/E AIR COOLER 1 INLET TEMPERATURE developed?

The questions were presented in a multiple-choice format, with the options of ”decreasing”, ”steady”, and ”increasing”. To calculate participants’ objective situation awareness, signal detection theory (SDT) is employed. Based on the two types of questions and two answer options, there were four possible outcomes: the parameter was changing and the participants detected it correctly (hit); the parameter was changing, but participant did not notice (miss); the participants correctly reported it as steady (correct acceptance); or the parameter was not changing, and the participants incorrectly reported it as either increasing or decreasing (false alarm). To assess the participants’ accuracy in evaluating the situation, sensitivity (d') as measurement of objectively is employed by using the below equation. Where Hit is the probability of hit occurring and FA is the probability of false alarm occurring. Sensitivity (d') is calculated for three categories: perception, comprehension, and overall situation awareness.

$$d' = z(p(Hit)) - (p(FA)) \tag{2.1}$$

Subjective situation awareness using the SART questionnaire, which was divided into three categories: attentional demand (D), attentional supply (S), and understanding of situation (U), based on the SA-constructed dimensions listed in Table 2.2. For each dimension, original 7-point Likert-type scales ranging from low to high, based on the dimension’s question. The perceived situation awareness was then calculated by eliminated the total understanding with the difference between attentional demand and attentional supply. From this point, a higher situation awareness indicated that attentional supply exceeds the attentional demand, and vice versa. The participants also completed the SART questionnaire after each trial; there are four sets of results for each participant.

$$SA = U - (D - S) \tag{2.2}$$

Table 2.2: SART dimensions and questions

Dimension	Question
Demands on attentional resources (D)	
Instability of situation	How changeable is the situation?
Variability of situation	How many variables are changing within the situation?
Complexity of situation	How complicated is the situation?
Supply of attentional resources (S)	
Arousal	How aroused are you in the situation?
Spare mental capacity	How much mental capacity do you spare in the situation?
Concentration	How much are you concentrating on the situation?
Division of attention	How much is your attention divided in the situation?
Understanding of situation (U)	
Information quantity	How much information have you gained about the situation?
Information quality	How good information have you been accessible and usable?
Familiarity	How familiar are you with the situation?

NASA-TLX and its six dimensions as a subjective measurement with 20-point Likert-type scale was used to evaluate workload subjectively. Previous research has demonstrated that NASA-TLX is a reliable measure of workload for a variety of work environments, including the engine plant simulator [39]. Perceived workload is calculated by adding up the dimensions in Tabel 2.3, after weighing them according to the participants' preference from multiple comparisons. After each trial, participants completed the questionnaire, resulting in four sets of results for each participant.

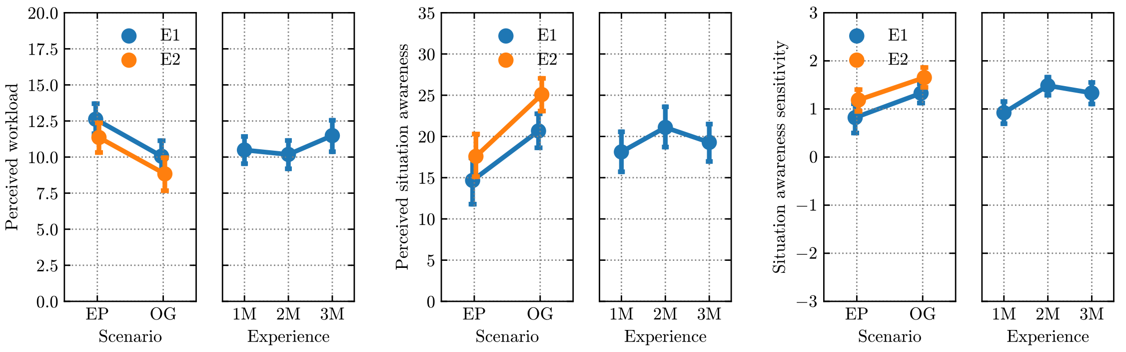
Table 2.3: NASA-TLX dimensions and questions

Dimension	Question
Mental Demand	How mentally demanding was the task?
Physical Demand	How physically demanding was the task?
Temporal Demand	How hurried or rushed was the pace of the task?
Performance	How successful were you in accomplishing what you were asked to?
Effort	How hard did you have to work to accomplish your level of performance?
Frustration	How insecure, discouraged, irritated, stressed, and annoyed were you?

2.4.3 Result and discussion

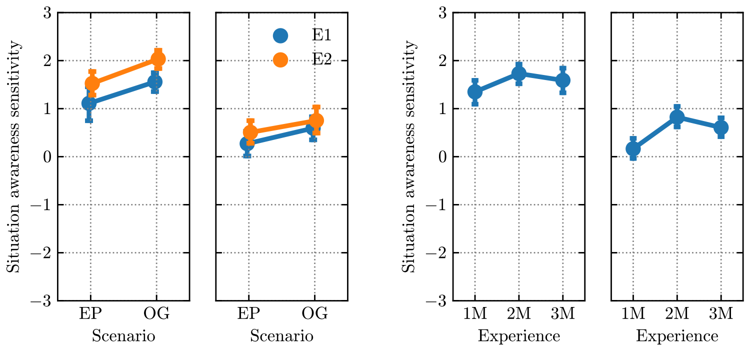
To analyze the result, several analytical methods such analysis of variance (ANOVA) is used. The null hypothesis then defined as there is no significant difference between variable measurement (dependent) under each level of independent variable. p-value less than .05 is use to neglect the null hypothesis. For the first experiment, since the participants was recruited based on their background on onboard training, the experience interference is separated from scenario level and participation level.

The result from perceived workload measurement is shown in Figure 2.4 (a). There is no two-way interference between the scenario and participation, tested with two-way ANOVA. However, participant reported higher workload in the entering port scenario, compare to the ocean-going scenario $(F(1,23) = 57.40, p < .01)$. This show designed scenarios where entering port scenario have high demand, were able to create two different workload It enables to verify the next analysis to examine the interference between situation awareness s and workload. In the effect of participation order, the participants reported higher workload for the firs-time participation, compare to the second participation $(F(1,23) = 11.47, p < .01)$. This confirmed the familiarity behavior of participant; the workload became lower when the human became used to the system. Separated one-way ANOVA for experience level effect show, there was no different among the three groups $(F(2,21) = 0.26, p > .05)$. This result reflects the subjective measurement limitation to discriminate the group experience level.



(a) (b) (c)

Figure 2.4: Measurement result for task-load and participation interference



(a) (b)

Figure 2.5: Measurement result divided to perception and comprehension

With the same pattern from workload analysis, Figure 2.4 (b) show the result for perceived situation awareness. There is no two-way interaction in perceived situation awareness effected by scenario and participation level. However, the participants reported higher situation awareness in the ocean-going scenario, compare to the entering port-scenario $(F(1,23) = 46.79, p < .01)$. This shows, in more demanding situations, the participants have lower perceived situation awareness. Thus, the correlation between workload and situation awareness can be explained. The second time participation give the participant perceived situation awareness value higher than the first time participation $(F(1,23) = 15.95, p < .01)$. This explains the situation awareness is involved when the human got familiar with the system. Similar to the perceived workload, one-way ANOVA shows there was no different from the experience effect. Similar to the subjective workload, subjective situation awareness is not sensitivity on the participants experience $(F(2,21) = 0.72, p > .05)$.

Figure 2.4 (c) show the interaction of scenario and participation level to the objective situation awareness (sensitivity). However, the participants have higher sensitivity during the ocean-going scenario compared to the entering port scenario $(F(1,23) = 39.03, p < .01)$. This reflects also when the situation awareness sensitivity break down into perception and comprehension as shown in Figure 2.5 (a). The different is both found in the perception level $(F(1,23) = 20.15, p < .01)$

and comprehension level $(F(1,23) = 15.21, p < .01)$. This explains the freeze-probe technique for objective follow the subjective measurement finding where the situation awareness tends to increase in less demanding situations.

For the effect of participation level, participants have higher situation awareness sensitivity in the second-time participation $(F(1,23) = 8.65, p < .01)$. This pattern followed in the perception $(F(1,23) = 9.94, p < .01)$ but not in comprehension level, but not with the participation $(F(1,23) = 3.30, p = .08)$. In the participation effect, the second time or more familiar condition gives the higher situation awareness, similar pattern with the subjective measurement. However, when it breaks to the perception and comprehension, this familiarity effect did not happen to increase situation awareness in comprehension level.

For the grouping based on the experience, one-way ANOVA shows the difference between the three groups $(F(2,21) = 4.80, p = .02)$. Post-hoc analysis done by Tukey-HSD bring the results of SA sensitivity of group of one-month experience has lower situation awareness compared to the two-month and three-month experience. Shown in the 2.5 (b) This finding is not found in the perception level $(F(2,21) = 1.93, p > .05)$, but it is exist in the comprehension level $(F(2,21) = 5.7, p = .01)$ with the same finding where the one-month experience has lower situation awareness compare to the other groups. The background of training contains can explained this result. The two-months and three-months training have a greater portion of engine watchkeeping.

2.5 Head-worn display utilization

The implementation of head-worn displays in work condition presents various advantages; however, it is crucial to investigate the potential drawbacks in the design and installation. The second experiment study aims to explore the effect of head-worn display application and how they can enhance department work, particularly in managing false alarms. Beside the workload and situation awareness that have been examined in the first experiment, there is trust in automation that is included in the second experiment.

Twelve marine engineering students, both undergraduate and graduate, with an average age of 21.7 $\pm$1.1 has participated as subjects. The participants were cadets who had been onboard a ship for duration of one month to one year. Based on their training and academic background, it was assumed that all participants were equally capable of performing the designated tasks during the experiments. informed consent was obtained from all participants through a participant agreement form. The recruitment of participants and experimental procedures followed an ethical code by the faculty board.

For the experimental design, *2x2* within participant design is employed. All participants were exposed to two information display status condition: on and off, while wearing head-worn displays continuously in both conditions to reduce additional workload interference. Additionally, participants were exposed to two levels of task load condition: high and low-load conditions, as indicated by the number of alarms activated in one session.

Participants visited the engine room simulator five times, three times during the training session and two times for data measurement. The training sessions allowed participants to become familiar with the experimental setup and tasks. The first data-measurement session was divided into two trials with different workloads, but with one head-worn display status. The second data-measurement session followed the same pattern but had different head-worn display status condition. The high and low-load scenarios each had 12 and 6 alarms introduced, respectively, and the ratios of true and false alarm were identical between the two scenarios. In the high-load scenario, there were 4 true and 8 false alarms, while the low-load scenario had 2 true and 4 false alarms.

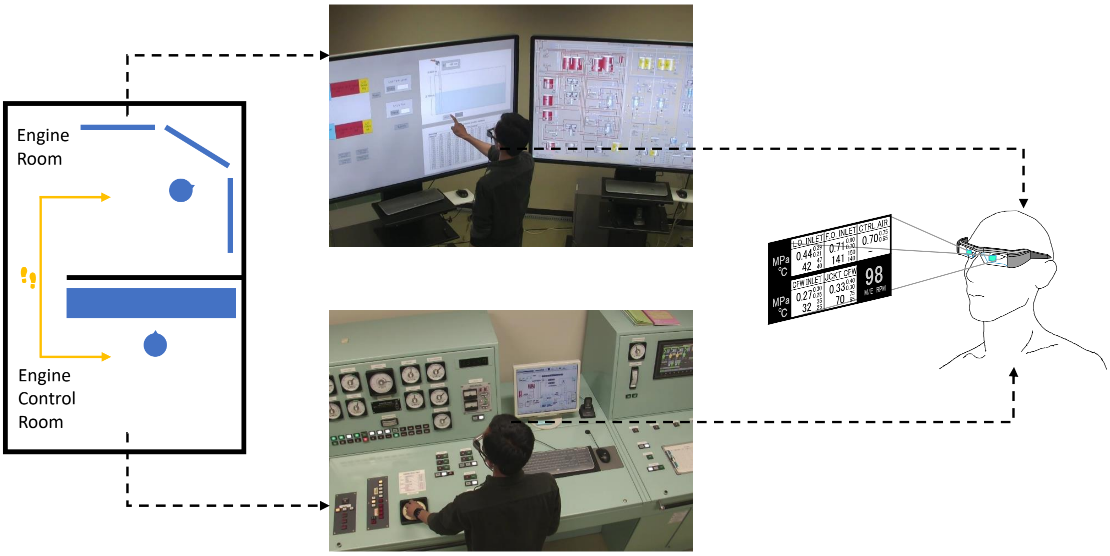


Figure 2.6: Engine plant simulator layout, ECR and ER side

2.5.1 Experimetal setup

The same engine room simulator was utilized to conduct the experiment. This time, engine room (ER) side is also used as shown in Figure 2.6. The ER was designed to simulate the equipment and components of an actual engine plant on a ship and featured two monitoring displays. Participants were able to control the system process, such as opening or closing valve, starting or stopping a pump, and performing other machinery operations. Similar with the first experiment, an ECC was installed in the ECR to monitor engine operation conditions and alarm indicators, allowing participants to acknowledge alarms and take necessary action to return to normal condition.

To replicate work on a ship, the participants were assigned two tasks. Firstly, they had to carry out maintenance tasks on the ECR side, such as opening or closing a valve, starting or stopping a pump, and conducting a tank-sounding task. These tasks required cognitive and spatial task-load capabilities as the participants had to locate the components along the flow line of the engine system on the monitor displays. Additionally, in some tasks like taking a tank-sounding measurement, the participants also had to conduct simple calculations to estimate the tank level. Secondly, the participants had to conduct a monitoring task by acknowledging the alarm from the ECR on its activation.

During the experiment, two alarm types were used: true and false alarms. True alarms were triggered when the process value exceeded (or fell below) a particular threshold due to engine operation conditions. False alarms, on the other hand, were triggered by sensor malfunction or ship movement and did not require any action. In each session, there are 12 and 6 alarms in the high and low-load situation, respectively, resulting in the same true and false alarm ratios. Specifically, there were 4 true alarms, and 8 false alarms in the high-load scenario: 2 true alarms and 4 false alarms in the low-load scenario.

The Epson Moverio BT-200, a binocular head-worn display shown in Figure 2.6, was used in this experiment. The binocular display was preferred because it's more comfortable and easier to focus on, as it can be observed by both eyes [40]. Through the experiment, participants were able to view the engine indicators on the ECC located at the ECR and on the head-worn display during the information-on scenario, as illustrated in left part of Figure 2.6. A subject matter expert (SME) selected the engine indicator parameters based on their importance during engine supervisory control, which included the process value, high-alarm threshold, and low-alarm threshold. An active alarm has different indicator colors on the display, and it is accompanied by audible alarms. All alarms were configured to be turned off automatically after 20 seconds one it has acknowledged.

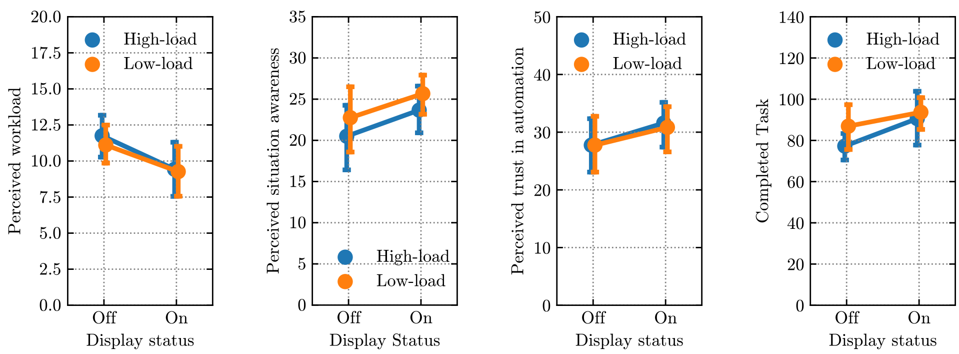
2.5.2 Measurement instrument

The measuring instrument in this scenario is following the first study to use the subjective and objective measurement. NASA-TLX is used to measure the task workloads in each trial. SART is used to measure the situation awareness subjectively. For trust in automation, is measured using TiA (Trust in Automation) questionnaire [41]. TiA has 12 dimensions, divided into the group of seven dimension measuring trust, and five dimension measuring distrust.

To measure objective performance, the simulator system recorded the response time of each participant. Response time (RT) was defined as the duration between alarm activation and deactivation. If a participant did not respond to an alarm, it was automatically deactivated, resulting in a maximum response time of 20 seconds. The simulator system also recorded the percentage of maintenance tasks completed by each participant.

2.5.3 Result and discussion

The analysis follow the same procedure as the first experiment, the result shown in Figure 2.7 (a). There is no two way interaction between display status and task-load ($F(1,11) = 0.04, p = .85$). For the perceived workload, participants reported that they have higher workload where the display status on the head-worn display is off ($F(1,11) = 7.96, p < .05$). This indicated when the participants have adequate information to discriminate between true and false alarm, they can obey to return to ECR. Therefore, the workload is lower since physical movement can be minimized. However, the number of activated alarm to make different task-load did not appear to make different perceived workload ($F(1,11) = 1.34, p = .27$). It turns the disability to analyze further the correlation between workload and situation awareness with trust in automation. From the verbal interview with several participants also found that they did not see any difference between high-load and low-load.



(a) (b) (c) (d)

Figure 2.7: Measurement result from display status and task-load interference

The perceived situation awareness is shown in Figure 2.7 (b). There is no two-way interaction between display stats on the display status and task-load ($F(1,11) = 0.04, p = .85$). In the single effect, there was only different in task-load interference ($F(1,11) = 6.97, p < .05$), but not in the display status ($F(1,11) = 2.58, p = .14$). This result is beyond the hypothesis since putting information on the head-worn display is aimed to improve situation awareness. In the perceived situation awareness, there may be a trade-off between attention demand and supply, where putting information will increase attention supply, but in other way also increase the attention demand. Therefore, putting continuous information on the head-worn display is not effective.

The perceived trust in automation result is shown in Figure 2.7 (c). There is no two-way interference between display status and task-load ($F(1,11) = 0.55, p = .47$). In individual effect, there is no effect by task-load ($F(1,11) = 0.33, p = .58$), but display status give the effect where

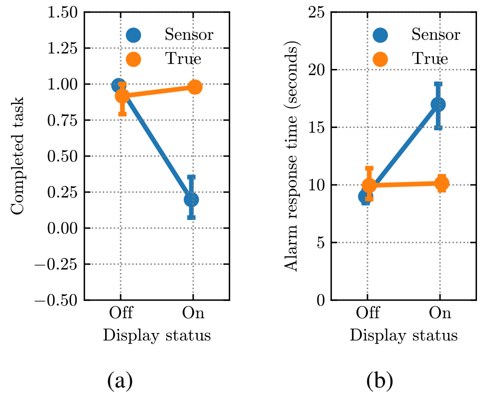


Figure 2.8: Measurement result from display status and alarm type

participants perceived more trust in automation during display status on, compare to display status off $(F(1,11) = 7.32, p < .05)$. This finding is then compared to the participant behavior in the objective measurement.

For the first objective measurement, the total alarm response ratio shown there was a two-way interaction between display status and alarm type $(F(1,11) = 41.18, p < .01)$, as shown in Figure 2.7. The following post-hoc analysis using T-test shows the participants respond less often to the false alarm during the display status on. This reflects the participants were hesitant to return to ECR to attend false alarm when the display status on. There is also tendency the participant follows "better-safe-than-sorry" approach to make a balance between maintenance task and monitoring task during display status off. This is shown as the participants take safe strategy; they return to the ECR every alarm sounded.

On the other objective measurement, alarm response time shown in Figure 2.8 (b).There is no two-way interaction between the variable $(F(1,11) = 0.51, p = .49)$. However, the participants take longer time to the response alarm in the session with display status on $(F(1,11) = 7.69, p < .05)$. Additional interview after the experiment shows the participants said they only look at the head-worn display information after the alarm sounded. Therefore, they took time to confirm this information and made the response time longer, compared to scenario where display off. Thus, there is trade-off between suitable response action to the response time in this case.

For the interaction between display status and task-load with completed task, is shown in Figure 2.7 (d). There is no two-way interference $(F(1,11) = 0.38, p = .55)$. However there is single effect from display status where the participant completed more maintenance task in display status on, but with less significance value $(F(1,11) = 12.97, p = .07)$. The total maintenance task completed is increased since the participant can safely neglect the false alarm. Thus, with trust in automation, they can prioritize multiple task efficiently.

2.6 Discussion

As the non-technical skills endorsed by IMO, situation awareness is well studied but only in the scope of navigational tasks. However, the work at the engine room operation with its ERM is different with strategy of BRM on the bridge. To address this gap, two experiments were conducted using an engine plant simulator to investigate situation awareness in the scope of human cognitive levels. Additionally, workload and trust in automation were also examined when developing a head-worn display as a cognition aid. The first experiment utilized a full-mission engine plant simulator to assess the feasibility and sensitivity of subjective and objective measurement methods for situation awareness, taking into account factors such as task load, familiarity, and experience. In the second experiment, the use of head-worn displays was investigated using the same simulator setup, revealing the balance between information availability, visual complexity, workload, and trust in automation when building situational awareness.

The sensitivity of subjective and objective situation awareness measurement methods varies. Although subjective measurement is a convenient method, it only responds to task complexity and not experience. On the other hand, objective measurement is sensitive to both task complexity and the participant's experience. The freeze-probe technique is a general-purpose objective measurement method that enables researchers or evaluators to develop scenarios relevant to the study's purpose and tailor the technique to the training's objective. Measuring situation awareness is crucial to non-technical skill training in the engine supervisory work, aiding in attention, contingency, and planning. The experimental results demonstrated that higher workload reduced situation awareness, highlighting the need for efficient information integration through established mental models and schemata. Additionally, measuring situation awareness can evaluate the process of designing interfaces for novices and assist with developing mechanisms for switching between goal-driven and data-driven tasks. Situation awareness oriented training and device design should prioritize this type of mechanism.

The benefit of using the head-worn displays in the engine room work is the reduction of engineers' movement between engine and control rooms, which lowers their workload. However, continuously displaying information on the head-worn display does not necessarily improve situational awareness because it can overload the visual attention. It is recommended to turn off the head-worn display when engine parameters are normal and activate it when the engine parameters are move towards alarm thresholds. The increased trust in automation is reflected in the alarm response ratio, where participants did not respond to false alarms when information was available on the head-worn display. This led to an improvement in maintenance tasks, allowing for safe neglect of false alarms from the control room. Participants prioritized multiple tasks efficiently with the information available, but response time was longer due to the confirmation process. Nonetheless, a suitable response action for engine control is more important than a short response time, and the trade-off between confirmation and response time can be accepted for better performance.

Although utilization of the simulator can be able to explore situation awareness at a cognitive level, there are several limitations to be considered. First is the participants were cadet student when the experiment is conducted. Therefore, the designed scenarios were simplified and limited representation, and several details were removed to comply with it. In the first experiments, there were no mitigation actions necessary to bring the system return to normal condition when the trouble or alarm is activated. In result, the mental model of advance troubleshooting was not investigated in this study. Future study may consider covering active action scenarios and also communication, since both of these factors are essential for critical situation management. One should consider also covering wider difference of cadet experience and to utilize more representative scenarios. Second is the number of participants is to be considered small. Therefore, qualitative study may be considered to be additional instrument regarding the number of participants. The small number of participants in the second experiment also made the experiment design to put participant to perform all scenario combinations. Thus, carry-over effect from participation to the next participation may exist and the objective performance data can be affected. Such kind of limitation can be overcome by inviting more participants and divided into between-group experiment design. The limitation also comes from the hardware side. Since the experiment used the engine plant simulator as the instruments, the condition became visual saturated at the second experiment. Thus, one may consider increasing the ability of head-worn display feature to be activated only when the alarm is activated.

Within these two experiments, the ability of human factors concept such as workload, situation awareness and trust in automation as demonstrated in this study is advantages not only to explore different states of human behavior but also to as evaluation methods of hardware or cogition aid to the work environment. However, several limitations are mentioned as before. In the next consideration, measurement of situation awareness should be addressed beyond individual human cognition and thus able to able to cover situation awareness effectiveness in team communication in engine department. Moreover, situation awareness should also studied between the engine room department and deck department, where the recent studi often treat both as separate work enviro-ment and conduct the study separately. Future studies in human cognition level may consider te actual engine room situation where evaluation such additional cognition aid still persist. There are lot of opportunities to examine the human cognition level.

Chapter 3

Situation awareness in human reliability analysis

3.1 Effort in quantify human reliability

Human error is the term that is often mentioned in maritime safety study to describe the human contribution in activated of an accident, regardless the human role also to mitigate the accident to occurs [3]. Human error can be breakdown into several parts and one of them, situation awareness loss, is defined as the disability of operator to maintain their situation awareness during the operations [21]. The initial error in situation awareness will result in the operator failure to understand the operation condition and it leads to the failure to take appropriate action. In maritime operation alone, the report by Grech et. al revealed 71% of incidents in maritime operation are due to the situation awareness loss [4]. This may be in line with the IMO action to include situation awareness as one of non-technical skill in crew resources management [42, 43].

Effort in quantifying the human error, or in contrary say, human reliability analysis (HRA), have been conducted in decades for various field [44]. In the maritime operation, IMO includes the HRA as one of the scopes in their Formal Safety Assessment guideline [45]. Based on the guideline, the HRA is treated the same as the technical system analysis, starting from hazard analysis, risk analysis and risk control option. In terms of methodology, the process of HRA and its quantification aligns closely with established risk analysis techniques such as fault-tree analysis and the utilization of historical data for risk quantification. The guideline also emphasized that the result from HRA is able to fit into hardware analysis.

$$HEP = \frac{Number\ of\ human\ error\ that\ have\ occured}{Number\ of\ opportunities\ for\ human\ error} \tag{3.1}$$

IMO defines human error as the condition that departure from acceptable or desirable practice in individual or group, resulted in unacceptable results [45]. Human error probability as explained in equation above, defined as the ratio of number of human errors compare to the number of opportunities for the human error itself occurs. Some human error quantification is already mentioned in IMO FSA guideline but most of the method use the bottom approach where various data is used, but commonly from the accident report. Compared to hardware reliability, accident reports are relatively few. This most of the time tackle by including another data [46], such as the expert judgment. Expert judgement can provide valuable insights into the factor but also can be subjective and biased. Advances in technology have enabled the use of simulations to evaluate the impact of human factors on system performance.

The development of HRA in maritime operation has contributed to several methods. Yang et al. for instance, modified the cognitive reliability and error analysis method (CREAM) by applying fuzzy and Bayesian inference to suggest instant estimation of human reliability performance for specific case in marine engineers' behavior [47]. The approach using combination of system theoretic process analysis (STPA) and success likelihood index methods (SLIM) give more perfective from the human-machine interaction system for LNG ship-to-ship bunkering [48]. For future projection, HRA in operation error for maritime autonomous has been conducted using combination of SLIM and fuzzy sets to predict error in maritime autonomous surface ships (MASS) operation [49].

Most HRA methods tread performance shaping factor (PSF) independently. Although they already recognized that the PSF undoubtedly contains some overlap [50]. PSF is defined as factors which have influence either is positive or negative effect on the human performance [45]. It supported by finding in the psychology and human factors to indicate that there is inter-relationship between PSF. The second challenge is using human error data alone also has a drawback since once can extract information on the frequencies of failure, but not number of successful performances [51]. Indeed, this probability is needed to failure probability empirically, but extraction from accident report alone should be expanded to include another perspective.

To answer the issue of dependency between PSF, the utilization of Bayesian network is steadily increasing in HRA methods [44]. By allowing for the modeling of complex dependencies and uncertainties, Bayesian networks offer powerful tools for analyzing it. Bayesian networks allow to model PSFs and their interaction in a probabilistic framework. Thereby, analysts assess the likelihood of human error and identify the most critical factors that contribute to it. The second advantage using Bayesian network comes from combining the data, since experts have big uncertainty. That is why we need to combine it with another data, to reduce uncertainty. On year-to-year analysis, utilization of the Bayesian network is starting to commonly use also in offshore industry [52] [53], ship engine operation [47], offshore emergency situation [54], and ship collision [55]. It has a wide application range from operator safety assessment to the evacuation training development.

The challenge in quantifying human reliability analysis is in integrating and combining information from diverse sources, including databases, simulator data, and expert judgement. Specifically, it is in characterizing the information and understanding how the parameters of Bayesian network interact with the newer information. This process enables researchers and analysts to make more informed decisions and predictions based on the combined evidence from multiple sources, enhancing the reliability and effectiveness of the modeling approach.

Situation awareness is not a general performance shaping factor, but it is a more popular human performance measurement in the field of human factor engineering and psychology. Specifically in study to evaluate digital control rooms where operator plays more important role in the field of human cognitive behavior [56]. This study aims to examine situation awareness as one of performance shaping factors in shaping human reliability analysis. The Bayesian network is used as modeling tools and the data from the experimental result is utilized to combine with expert judgement.

3.2 Bayesian network for model construction

Bayesian theorem has different differences for model construction and updating. In the model construction, Bayesian inference is used, while in the model updating, Bayesian inference is applied. Bayesian network, where the graphical method is used does not have to follow the Bayesian inference. It is called Bayesian network since has similar rules for inference.

In Bayesian network, a causal model is established through directed acyclic graphs composed of nodes and arcs, with nodes serving as variables and arcs representing the causal link between the nodes. The target of arcs called child nodes, where the source of arcs called parent nodes. The parent nodes that act as the initial node in graph without another parent node are called root nodes. The impact of one node to another node, if it discreate, can be quantitatively represented by a conditional probability table (CPT). The value of CPT then shows the influence of parent nodes.

The stage in model construction using the Bayesian network is shown in Figure 3.1. Here,

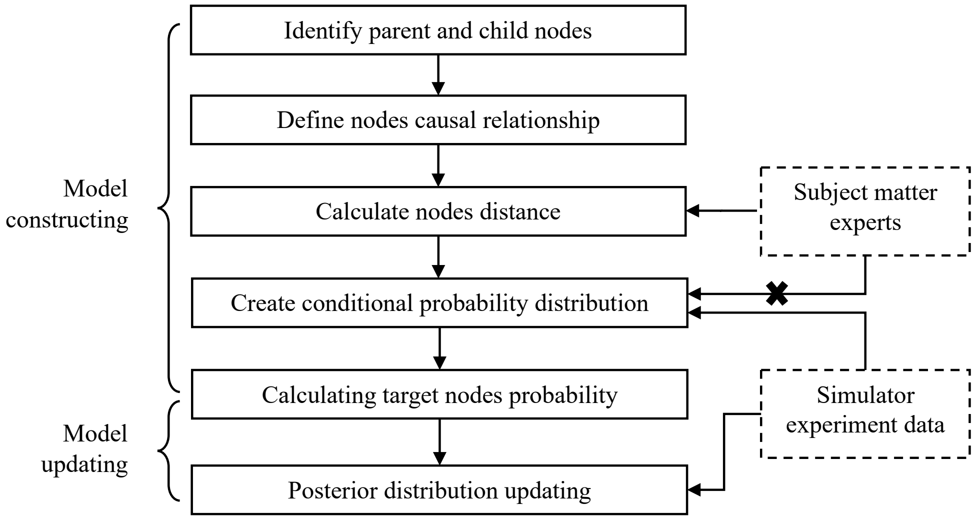


Figure 3.1: Proposed flow of model construction and updating

the recent model in HRA for maritime operation is applied a big portion of its decision on the subject matter expert. Specifically in Bayesian network, the subject matter experts are often used to weighing the importance of each node to calculate the node distance and give the input directly in creating the conditional probability distribution. In this study, the role of subject matter expert is still important, but to reduce the additional uncertainty based on the lack of knowledge and bias, the scope of subject matter experts in creating the conditional probability distribution is eliminated. As a substitution, the simulator experiment data is employed. The detail of integrating the simulator experiment data with the Bayesian model will be explained below.

3.2.1 Experiment based construction

An advantage of construction of Bayesian network is its capability to construct causal framework. As shown in 3.2, the proposed model employs a causal factor mapping to estimate the probability of human failure event (HFE) as the target of the analysis. To construct this, there are three levels of construction as shown in Figure 3.2 in general, and Figure 3.3 in specific from the simulator experiment design from Chapter 2.

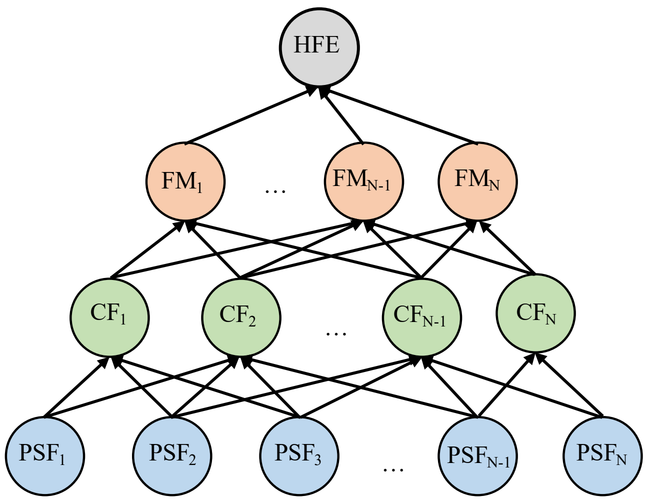


Figure 3.2: General structure of proposed Bayesian network

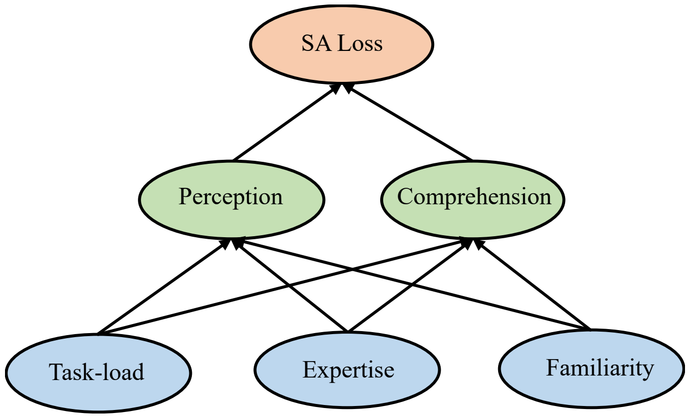


Figure 3.3: Proposed Bayesian network based on the experiment

The state of failure mode (FM) as the first stage below the human error, has a function to determine the probability of human failure event. This stage also explains the various types of supporting events leading to HFE. In this case, the loss of SA is assigned as a single FM contributing to HFE. Although it only covers the initial process of information processing, other factors such as decision making, and response execution may be considered as another FM nodes.

The cognitive function (CF) layers lie directly below the FM in the proposed model and consist of variables that closely influence FM. CF is the variable that can be observed in the controllable environment such as simulator but not in the real environment. In the experimental design, this is what interpret as the dependent variable. Based on the first experiment in Chapter 2, two CF nodes of perception and comprehension. As the definition from Chapter 2, perception is how the operator consumes the cues from the experiments, comprehension is stage where the information is combined and make meaningful representation.

Performance Shaping Factors (PSF) represent the final layer of the proposed model and are considered the nodes that can be observed in a real workplace situation. PSF are constructed based on the independent variables in the experiments, with three specific factors: task-load, expertise, and familiarity. These factors are impacting the situation awareness in the stage of perception and comprehension. Task-load refers to the amount and complexity of the work being performed, while expertise reflects the level of skill and knowledge possessed by the worker. Familiarity relates to the worker's experience with the task or environment, which can affect their ability to recognize and respond to potential hazards.

3.2.2 Node relative importance

Since the child nodes have multiple parent nodes, each parent node in influencing the child nodes is represented by relative importance. The best practice to determine it is by using the weight W_i for each parent node i [52]. In this study, AHP is using pairwise comparison is used. The list of academia who also have professional onboard experience as chief engineer, listed in Table 3.1. They gave the weighting by questionnaire that provided individually after interviewing.

Table 3.1: Experts profile

Expert	Professionals	Onboard (years)
1	Academics	15
2	Academics	5
3	Academics	14.5
4	Academics	10

The result from the weighting process expresses the relative importance of workload, expertise and familiarity are $W_T = 0.26$, $W_E = 0.41$, $W_E = 0.33$, respectively. While in the construction of comprehension, three PSF have relative importance of $W_T = 0.13$, $W_E = 0.55$, and familiarity with $W_F = 0.32$. For the relative importance of CF node to construct FM node, the perception has $W_P = 0.14$ and comprehension has $W_C = 0.86$.

3.2.3 Node distance

The relative importance result obtained from the weighting process was utilized to establish the relationship between child nodes and their respective parent nodes. In other ways, the distance between the child node and its parent node was defined based on their relative importance. This approach was employed to assign lower probability values to the child node when its parent nodes were in different states. Given the example, if the PSF node task-load were high, expertise is in high state, and familiarity is in good state, the CF node perception should have higher probability of being in the distance compared to medium and low state.

$$D_j = \left| \sum_{i=j}^{n} D_{ij} W_i \right| \qquad D_j \in [0,1] \tag{3.2}$$

3.2.4 Conditional probability for child nodes

$$P_j = \frac{e^{-RD_j}}{\sum_{j=a}^{c} e^{-RD_j}} \qquad P_j \in [0,1] \tag{3.3}$$

The equation above is used to quantify the distance between the child node and its parent nodes. It can confirm the condition, for instance, the PSD node task-load is in the easy state, expertise in high state and familiarity in good state, the child node perception should have higher probability of being in the high state compared to other state. In other says, this distance method assigns smaller probability if the parent nodes are in different states. Different with Røed, Li method to consider the direction of change of its parent nodes into absolute value is used in the question above [52] [56].

Table 3.2: Perception and comprehension sensitivity (d) from simulator experiment result

Variable	**Perception, comprehension**							
Task Load	Easy				Complex			
Familiarity	Good		Bad		Good		Bad	
P1	1.40	0.54	0.42	0.42	1.81	1.38	0.95	0.00
P2	1.38	0.42	1.40	0.95	1.94	0.42	0.00	0.00
P3	2.35	1.38	1.38	0.42	1.40	0.95	-0.42	0.00
P4	2.77	0.95	0.95	0.42	2.35	0.00	0.86	0.00
P5	2.35	0.42	1.38	1.40	1.40	0.54	0.97	0.42
P6	2.77	1.94	1.93	0.00	2.77	0.42	2.77	0.54
P7	2.35	0.95	1.38	0.95	0.97	1.81	1.94	-0.54
P8	2.77	0.42	1.81	0.95	0.97	0.00	1.40	0.95
P9	1.38	1.38	1.93	0.95	1.93	1.81	2.35	0.97
P10	1.40	0.54	0.95	0.42	1.81	0.00	-0.54	0.00
P11	1.38	0.00	1.81	0.00	0.97	0.00	1.38	0.00
P12	1.40	0.42	1.38	0.95	1.81	0.42	0.54	0.00
P13	1.81	-0.54	0.97	-0.42	0.97	-0.42	-0.43	0.00
P14	2.77	0.00	0.95	-0.42	0.54	0.54	0.97	0.00
P15	1.94	0.00	1.94	-0.42	0.95	-0.42	0.86	0.00
P16	2.35	0.42	1.38	0.54	2.35	0.42	0.97	-0.95

Table 3.3: Threshold and number of trial grouping

State	Perception		Comprehension	
	Threshold	Trial	Threshold	Trial
Low	$d \leq 0.86$	9	$d \leq 0.00$	27
Mediate	$0.86 < d < 2.35$	42	$0.00 < d < 1.4$	33
High	$d \geq 2.35$	13	$d \geq 1.40$	4

Table 3.4: Experiment result distribution for perception

Parent or CF		**Pr (CF | Parent State)**							
Task-load		Easy				Complex			
Experience		High		Low		High		Low	
Familiarity		Good	Bad	Good	Bad	Good	Bad	Good	Bad
Perception	High	0.75	0.00	0.25	0.00	0.25	0.13	0.13	0.13
	Mediate	0.25	0.88	0.75	1.00	0.75	0.50	0.75	0.38
	Low	0.00	0.13	0.00	0.00	0.00	0.38	0.13	0.50

The Table 3.2 shows the situation perception and comprehension sensitivity from the experiment in Chapter 2. The high value means higher situation awareness: the participants are more able to discriminate the question in freeze-probe. Based on the participant recruitment process, P1 until P8 are categorized with expertise high, while P9 until P16 categorized expertise low. Two thresholds are then applied to divide the sensitivity categorized into three states to comply with the model explanation: low, mediate, and high. Applying the threshold must be comply to the best practice that remains the average distribution, which in this case is the mediate state. Table 3.3 shows two thresholds that are defined by adjusting the number of *hit* and *false alarm* in situation awareness freeze-probe calculation. By defined these two threshold, the sensitivity (d) in Table 3.2 can be grouped and the number is shown as number of trials that included in low, mediate and high state.

In the previous paper, R value is decided from the expert judgment [52]. The variation assumed the R to be 1 to 5, is compared to the conditional probability retrieved from the conditional prob-

Table 3.5: Experiment result distribution for comprehension

Parent or CF		**Pr (CF | Parent State)**							
Task-load		Easy				Complex			
Experience		High		Low		High		Low	
Familiarity		Good	Bad	Good	Bad	Good	Bad	Good	Bad
Comprehension	High	0.13	0.13	0.00	0.00	0.13	0.00	0.13	0.00
	Mediate	0.88	0.75	0.50	0.50	0.63	0.38	0.38	0.13
	Low	0.00	0.13	0.50	0.50	0.25	0.63	0.50	0.88

Table 3.6: Conditional probability distribution for perception

Parent or CF		**Pr (CF | Parent State)**							
Task-load		Easy				Complex			
Experience		High		Low		High		Low	
Familiarity		Good	Bad	Good	Bad	Good	Bad	Good	Bad
Perception	High	0.80	0.34	0.23	0.10	0.44	0.13	0.11	0.03
	Mediate	0.17	0.55	0.64	0.46	0.46	0.64	0.55	0.17
	Low	0.03	0.11	0.13	0.44	0.10	0.23	0.34	0.80

Table 3.7: Conditional probability distribution for comprehension

Parent or CF		**Pr (CF | Parent State)**							
Task-load		Easy				Complex			
Experience		High		Low		High		Low	
Familiarity		Good	Bad	Good	Bad	Good	Bad	Good	Bad
Comprehension	High	0.54	0.36	0.25	0.19	0.46	0.29	0.23	0.16
	Mediate	0.30	0.42	0.46	0.35	0.35	0.46	0.42	0.30
	Low	0.16	0.23	0.29	0.46	0.19	0.25	0.36	0.54

ability from experiment result. Below RMSE function is applied to find the lowest difference. It simply compares the probability from the experiment e and the probability from Bayesian model f in the scenario i and the state j, where the m and n explain the number of scenario and states respectively. Based on the calculation, R value then decided as $R = 3.15$ for the perception and $R = 1.22$ for comprehension. The conditional probability distribution (CPD) as the result is shown in Table 3.6 and 3.7.

$$RMSE = \sqrt{\frac{1}{mn}\sum_{i}^{m}\sum_{j=1}^{n}(e_{ij} - f_{ij})^2} \tag{3.4}$$

3.2.5 Calculating target node probability

An alignment is needed since the situation awareness loss as FM node is binary node with either true or false state, while the CF has three states. The failure mode is calculated based on the probability of the CF nodes with P_{ik} with state from $k = a, b, c$, while the weighted value W_i was retrieved same step as the previous step using subject-matter experts. Q_{ik} is the 100-fold for the P_{basis} adjustment, where P_{basis} is retrieved following the previous study that use $P_{basis} = 0.01$.

$$P_j = P_{basis}\sum_{i=1}^{n} W_i \sum_{k=a}^{f} P_{ik}Q_{ik} \tag{3.5}$$

To demonstrate how the scenario combination produced the probability of situation awareness loss, the example of scenario with experience in high state, familiarity in bad state, and task-load in complex state is calculated with below equation. The input is the probability distribution of perception with 0.13, 0.64, and 0.23 and the probability distribution of comprehension is 0.29, 0.46, and 0.25.

$$\begin{aligned} P_{fail} &= P_{basis}(0.14x(0.13x0.01 + 0.64x1 + 0.23x100) + 0.86x(0.29x0.01 + 0.46x1 + 0.25x100)) \\ &= 0.253 \end{aligned}$$

3.2.6 Validation

The validation of the Bayesian network is barely conducted in the correction level of its value. However, several steps can be followed such as using the comparison with the expert judgement as follows. In this section, the expert gives the free scale measurement of how likely the participant will fail in maintaining the situation awareness in each scenario. As the result, shown in Fig. 3.4, Bayesian network result's pattern is in the same direction as the expert expectation, except in the scenario where familiarity is good, and expertise is low. In this intersection, the Bayesian network result has tendence of higher probability.

Another validation step as sensitivity test is to satisfy several axioms [55]. One of the axioms states if a child node has more than one parent node, the combination of each parent nodes effect should always stronger than the sole effect of parent node separately. This satisfied by explanation shown in Table 3.8. The combination where only one parent node in its lowest stage gives the target nodes situation awareness failure probability half, compare if all parents' nodes are in its lowest stage.

Table 3.8: Combination of lowest state effect

Task-load	Expertise	Familiarity	SA failure probability
Complex	High	Good	0.18
Easy	*Low*	Good	0.27
Easy	High	*Bad*	0.22
Complex	*Low*	*Bad*	0.58

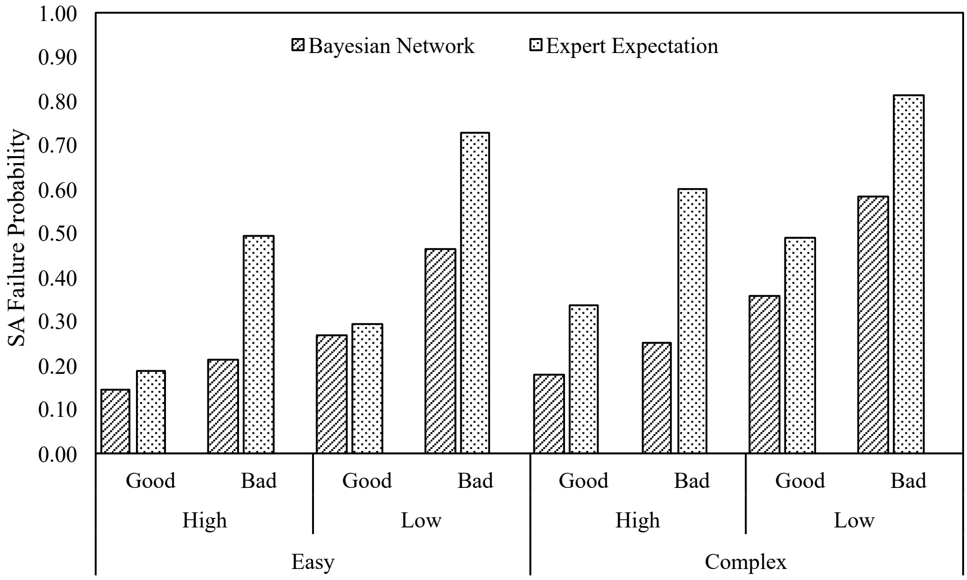


Figure 3.4: Comparison of result from Bayesian Network and Expert Expectation

3.3 Bayesian inteference for model updating

$$P(H|D) = P(H)\frac{P(D|H)}{P(D)} \tag{3.6}$$

The second stage after the model was constructed is model updating. In this section, the probability of each combination of performance-shaping factor, cognitive function, and failure mode is recalculated based on whether new experimental data from the simulator is available. These methods use Bayesian inference, as described in the equation above. To update the posterior distribution $P(H|D)$ using the prior knowledge distribution $P(H)$ and data collected from the likelihood mode $P(D|H)$ using the prior knowledge distribution $P(H)$, which is normalized with the distribution probability $P(D)$. From this perspective, the posterior data from one modeling process becomes the prior data for the next modeling process.

3.3.1 Construction of prior distribution

The binomial distribution, which is described by the equation below, can be used to explain how human error events can be attributed to situation awareness factors. The distribution represents

Table 3.9: Prior distribution p_0

Expertise	Task-load	Familiarity	$E(p_0)$	α	β	Prior Distribution
High	Easy	Good	0.145	0.5	2.9	$p_0, Beta(0.5, 2.9)$
		Bad	0.215	0.5	1.8	$p_0, Beta(0.5, 1.8)$
	Complex	Good	0.179	0.5	2.3	$p_0, Beta(0.5, 2.3)$
		Bad	0.253	0.5	1.5	$p_0, Beta(0.5, 1.5)$
Low	Easy	Good	0.269	0.5	1.4	$p_0, Beta(0.5, 1.4)$
		Bad	0.465	0.5	0.6	$p_0, Beta(0.5, 0.6)$
	Complex	Good	0.357	0.5	0.9	$p_0, Beta(0.5, 0.9)$
		Bad	0.584	0.5	0.4	$p_0, Beta(0.5, 0.4)$

the uncertainty around the number of failures (x) that occur in each set of demands (n), with the parameter of failure probability (p). This parameter is often uncertain and can be derived from either expert judgment or data. Groth et al. have suggested that Bayesian inference can be used to obtain this probability p [57].

$$f(x|p) = \binom{n}{x} p^x (1-p)^{\beta-1} \tag{3.7}$$

The Beta distribution was employed using a similar approach as the previous, to determine the p. The equation below represents the probability density function, which expresses the beta function $B(\alpha, \beta)$. From the explanation from Groth et al, the distribution can be conjugated with binomial distribution, where the posterior distribution parameters a_{post} and b_{post} are assigned using the equation below [57]. To proceed with Bayesian inference, the value of α and β need to be assigned.

$$f(p; \alpha; \beta) = \frac{p^{a-1}(1-p)^{\beta-1}}{B(\alpha, \beta)} \tag{3.8}$$

$$\alpha_{post} = \alpha_{prior} + x \tag{3.9}$$

$$\beta_{post} = \beta_{prior} + n - x \tag{3.10}$$

The contained non-information (CNI) distribution to define the prior distribution P_0 is utilized. The equation below depicts that the number of failures in the prior distribution is estimated as a, while the denominator $(\alpha + \beta)$ is interpreted as the number of demands. Following the existing study, the beta distribution in this case is constructed by a value 0.5, while b value is obtained using the equation below. Table 3.9 displays the extracted prior distribution P_0 from the experimental data.

$$E(Beta(\alpha, \beta)) = \frac{\alpha}{\alpha + \beta} = E(p_0) \tag{3.11}$$

3.3.2 Posterior distribution updating

Equation 3.8 and 3.9 is used to explain the process of updating the prior distribution with new experimental data. The prior distribution P_0 represented by its beta distribution $Beta((a_{prior}, \beta_{prior})$, is updated to obtain the posterior distribution P_1 with its distribution $Beta\,(a_{post}, \beta_{post})$.

Table 3.10 displays display with additional eight participants. The calculation of sensitivity is similar to that in Table 3.2 calculated using signal detection theory in freeze-probe method. To recall the previous definition, sensitivity here reflects the participant's ability to distinguish between hit and false alarm; higher sensitivity indicating better performance and zero sensitivity

Table 3.10: Experiment result to update the model

Variable	**Perception and comprehension**							
Task Load	Easy				Complex			
Familiarity	Good		Bad		Good		Bad	
P17	1.81	1.38	1.81	0.54	0.95	0.00	2.35	0.95
P18	1.81	0.42	1.40	0.00	0.43	0.00	0.95	0.00
P19	2.35	0.95	2.35	0.95	1.40	0.54	1.40	0.42
P20	1.94	0.54	1.81	0.54	1.40	0.00	2.35	0.54
P21	1.81	0.43	1.94	1.38	1.40	0.97	1.81	0.97
P22	2.35	1.38	2.35	2.35	1.40	0.97	1.38	1.38
P23	1.81	2.35	1.38	0.42	2.35	0.42	0.00	1.40
P24	2.35	1.38	2.35	0.95	2.35	1.38	1.81	0.42

Table 3.11: Posterior distribution p_1

Expertise	Task-load	Familiarity	Datax/n	α	β	Posterior Distribution	$E(p_1)$
High	Easy	Good	0/8	0.5	10.9	$p_1, Beta(0.5, 10.9)$	0.044
		Bad	1/8	1.5	8.8	$p_1, Beta(1.5, 8.8)$	0.145
	Complex	Good	3/8	3.5	7.3	$p_1, Beta(3.5, 7.3)$	0.324
		Bad	2/8	2.5	6.4	$p_1, Beta(2.5, 6.4)$	0.282

indicating inability to discriminate it. Further, in this updating section, the sensitivity of 0 is categorized as situation awareness (SA). Using this calculation, we assigned the value of $n = 8$ is the number of opportunities of failure and identified the number of failures x. Table 3.11 provides an explanation of the parameters distribution P_1.

3.4 Discussion

Bayesian network and Bayesian inference in general were demonstrated to constructed human reliability analysis, specifically for ship engine operation. Its effectiveness in combining the data from simulator experiment and expert judgment were used to construct and updating the human error probability. The advantage is given when the new data from simulator exists, the probability value is updated to the newest without reconstructing the model. Additionally, Bayesian networks demonstrate their strength in handling variable dependencies, which is crucial in accurately constructing human reliability analysis.

The construction of three stages: PSF, CF, FM to accommodate the experiment design is also able to explain the human factor construction in theoretical perspective. PSF, denoted as independent variable in experimental design is designed to cover the possibility to add more scenario construction in simulator data. Meanwhile, CF that covers dependent variable can rely on the measurement method in hand. And lastly FM in this case is demonstrated as a single variable of situation awareness loss. Thus, the possibility to expand it if the experimental setup has ability to cover more PSF, the measurement methods can cover more CF, and of course more FM instead of single situation awareness can be covered.

While the Bayesian network is a valuable tool for human reliability analysis, it contains certain limitations. First, the initial construct is often weak on detecting all PSF exists. Determining which PSF to include in the model becomes a trade-off, as it is challenging to capture the full range of factors that influence human performance. This limitation can potentially lead to an incomplete representation of the complex relationships between variables. Second, utilizing the data from the simulator data is not without a drawback since the participation during the experiment is conscious that their action will not have any consequences in the real system. Otherwise, knowing that their actions are being observed also motivates them to perform better than they would under normal circumstances. These limitations emphasize the need for caution and careful interpretation when using Bayesian networks in human reliability analysis.

Human error in general, that calculate by the ratio of number of human errors occurred and number of opportunities for human error is the approach methods came from the hardware reliability where the number of opportunities is defined as the duration of the component or part running over the time. In terms of human or operators' opportunities, the duration in the time is unlikely to be measured and recorded. The undefined divider in human error remains on each method. For example, in this study using the simulator data, the opportunity of error is translated into the number of trials in a sample data. Thus, to combine the human error result from one method to other methods need adjusting step to normalize the divider in human error probability.

Compared to the shore industry such as nuclear power plant where the general PSF is already defined and agreed as the standards to record the the human error probability, there is no exist the PSF in maritime operation. Perhaps since the area of work on board is also divided into two mains are of bridge and engine room, defining the PSF for general ship operation is not an easy task. The future may interact with this issue. The general PSF for maritime operation then can became reference under the IMO guideline, especially in formal safety assessment.

Chapter 4

Situation awareness in socio-technical system

4.1 Engine resources management between onboard and shore

Embark from the scope of internal human cognition and human-machine interface. The concept of situation awareness in maritime context is also considered to be important in the scope of the humans in complex environment of sociotechnical system. More specifically in the ERM, the balance between onboard team and shore support is critically argued to support safety operation [58]. The shifting also happens in the safety understanding lately. From the understanding of accident has special cause to the accidents as the product of relation in complex system [59, 60]. For thus, engine room resources management in the sociotechnical system where the engine room onboard and shore support is to be considered more in this study.

Ship operation has characteristic of distributed operation; not only onboard operation but also shore staff. It indicated the geographical operation is distributed in different places, and operation time also temporarily different, where most of them work in different shift or time difference [61]. As mentioned in Chapter 2 and 3, despite decreasing number of accidents in the maritime operations, human error especially the failure to maintain situation awareness is still dominant [4]. The lack of situation awareness is not only due to challenges in effectively communicating mental model among onboard team members but also between onboard team and the shore team [62]. Further, it is developed into individual perceptions of the overall situation that are based on incomplete and inaccurate information. In distribution team context, the team remain locked into a false understanding of situation until accident or incidents eventually occurs [21].

The finding from the alarm performance assessment for instance, implicates such practice to updating how the operator on board handle the situation is not barely a practice to be conducted in shipping industry. The alarm data remains onboard, and it is not mandatory to report the alarm rates or indication of other alarm performance assessment such as alarm floods or which one is considered to be catering alarm. There is a possible mismatch between the onboard operation and shore support management to update their situation awareness about the current operation. Therefore, to maintain situation awareness between two parties in sociotechnical system is the major point to in recent engine resources management.

To normally operated, there is often a trade-off to balance the efficiency and thoroughness in the ship operation [63]. It's natural for an organization such as shipping, ship owner, or ship management company to operate based on the optimum cost-effectiveness; the hidden layers bring the operation also to the accident [64]. The interaction of onboard operation and shore support needs to be deeply observed, but in the wide perspective of socio-technical system. Traditional approach on system safety is inefficient to tackle this issue. The study in this chapter aims to model the system dynamic to provide holistic interaction between safety and pressure allocation in sustaining safe operation.

4.2 Dynamics safety theory

The concept of safety in complex sociotechnical systems extends beyond the traditional view of the absence of errors and accidents. Safety is a dynamic property of the system. It encompasses proactive measures, risk management, and resilience. For thus, safety should be seen as system characteristic rather than an individual performance [65]. This perspective aligns with the notion that safety emerges from the interaction and interdependencies of various components within the system. As the importance of system think, safety can be seen as the property of the system, and it arises from the effective integration of human, organizational and technological system [2]. Therefore, understanding and enhancing safety in sociotechnical system necessitate a comprehensive approach that acknowledge the complex interaction and interrelationship between various elements as it aims to optimize the system performance while minimizing the risk and hazards [66].

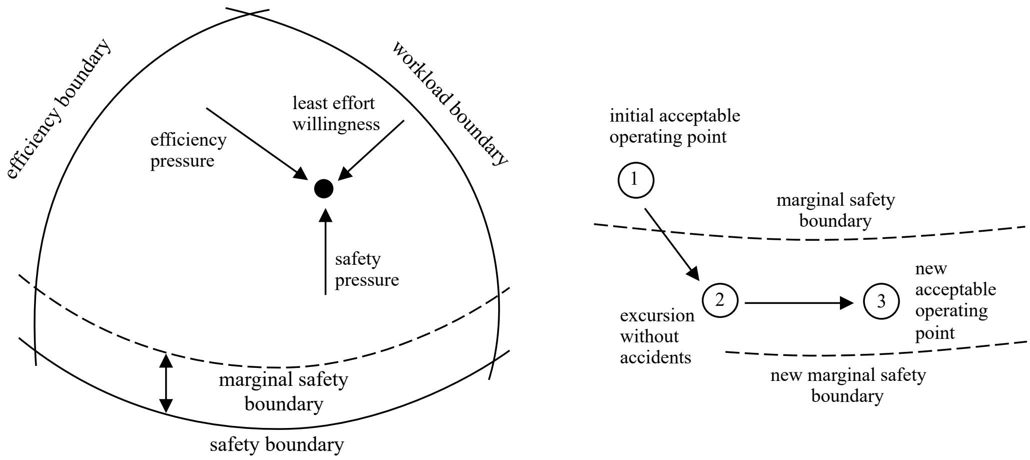


Figure 4.1: Rasmussen's dynamics safety model

Rasmussen with his theory introduces the descriptive side of safety dynamic in socio-technical system [64]. Illustrated from the Figure 4.1, an organization can operated on the feasible three boundary: efficiency boundary, safety boundary and workload boundary. The operating point is then influenced by three pressures according to their boundary. First pressure is to move far away from efficiency boundary to gain more profit. The second pressure is to move far away from safety pressure to avoid loss by incident or accidents. And the last pressure came from the field worker who, in daily routine, tries to reduce the workload. Since the system in considered to be dynamic, the operation point of the organization is continuously move and reduce the movement once the movement is small and random [67]. Once the organization passes the safety boundary, it means an accident. Since this boundary is difficult to define before the accident happens. There is another boundary called marginal safety boundary. This boundary marks the acceptable limit of accident result if the operation point stays within this marginal boundary. Crossing this marginal boundary what it called a near miss.

In Rasmussen theory, it is not only the operating point that considered move continuously but also this marginals safety boundary [64]. The location of it is the product of sociotechnical process. The occurrence of an accident will bring the marginal safety boundary far from the safety boundary. In other hands, long operation duration without an accident may bring the marginal boundary move outward closer to the safety boundary. There is also a testing movement, what Morrison defined as excursion pointed, as shown in Figure 4.1 left side, to deliberately cross the marginal safety boundary. As mentioned before, because this practice only produces near misses, does not immediately produce accident, the organization then adjust the marginal safety boundary lower. The organization remains in the operating point with new acceptable operating point.

4.3 System dynamics to model socio-technical system

In Rasmussen theory, it is not only the operating point that considered move continuously but also this marginals safety boundary [64]. The location of it is the product of sociotechnical process. The occurrence of an accident will bring the marginal safety boundary far from the safety boundary. In other hands, long operation duration without an accident may bring the marginal boundary move outward closer to the safety boundary. There is also a testing movement, what Morrison defined as excursion pointed, as shown in Figure 4.1 left side, to deliberately cross the marginal safety boundary. As mentioned before, because this practice only produces near misses, does not immediately produce accident, the organization then adjust the marginal safety boundary lower. The organization remains in the operating point with a new acceptable operating point.

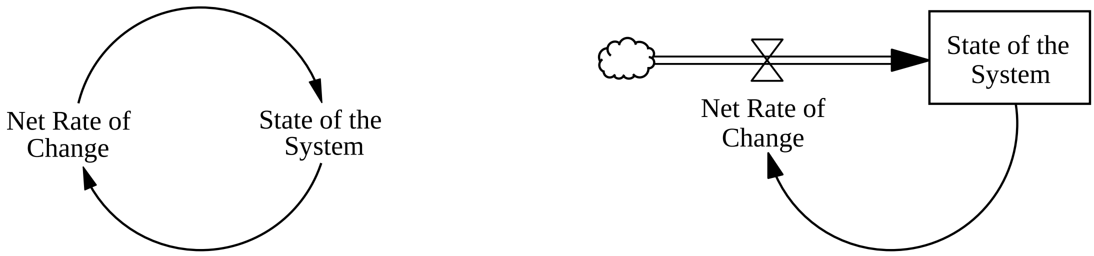


Figure 4.2: Causal-loop Diagram and Stock-and-flow

Construction of SD consists of quantitative and qualitative modelling. The qualitative part used the Causal-loop diagram (CLD) to map the causal relationship between the variables within a system. The CLD also draws the feedback loops to demonstrate the system behavior [68]. CLD enhances the understanding of models. While the quantitative part uses the stock-and-flow, equipped with differential equations. In this Chapter, the SD is used to model the dynamic safety theory as mentioned before, with application on ship operation.

A CLD is simply visual representation to describe the system on the system dynamics modelling. This diagram capture the interation between various components and highlight the feedback mechanims from the loops. The interconnected loops indicate how the system influece each others, either it is reinforce (positife feedback) or balancing (negative feedback). The arrow in the diagram indicates the direction of influence, how one variable change can lead to change in others, vice versa. Thus, CLD helps to understand how the systems behaves over a certain peiod, emphasizing the role of self-reinforce or self-correcting mechanism.

While CLD is used as qualititave method, stock-and-flow diagram are fundamental concept used for quantitative study purposes. Stocks are described as variables that accumulate quantities of interest within a system. Flow, that often describe like a valve, on the other hand, describe the rates of change in stocks over the time. It indicates the inflow or outflow of the quantity from or into a stock. Stock accumulated of te net inflow and outflow through the associated flows and store the history of the system behaviour and represent the current state of the variables. The understanding of quantitaive approach allow for the analyis can be conducted to observe long-term behaviour.

4.4 Causal loop diagram construction

The base model in this CLD development is based on the Rasmussen safety dynamic theory that first approach to model it was done by Morrison and Wears [60]. In their approach, workload and safety was simplified as a single boundary. In maritime operations where the team is divided geographically and temporally, such kind of unique approach is needed. The CLD was developed

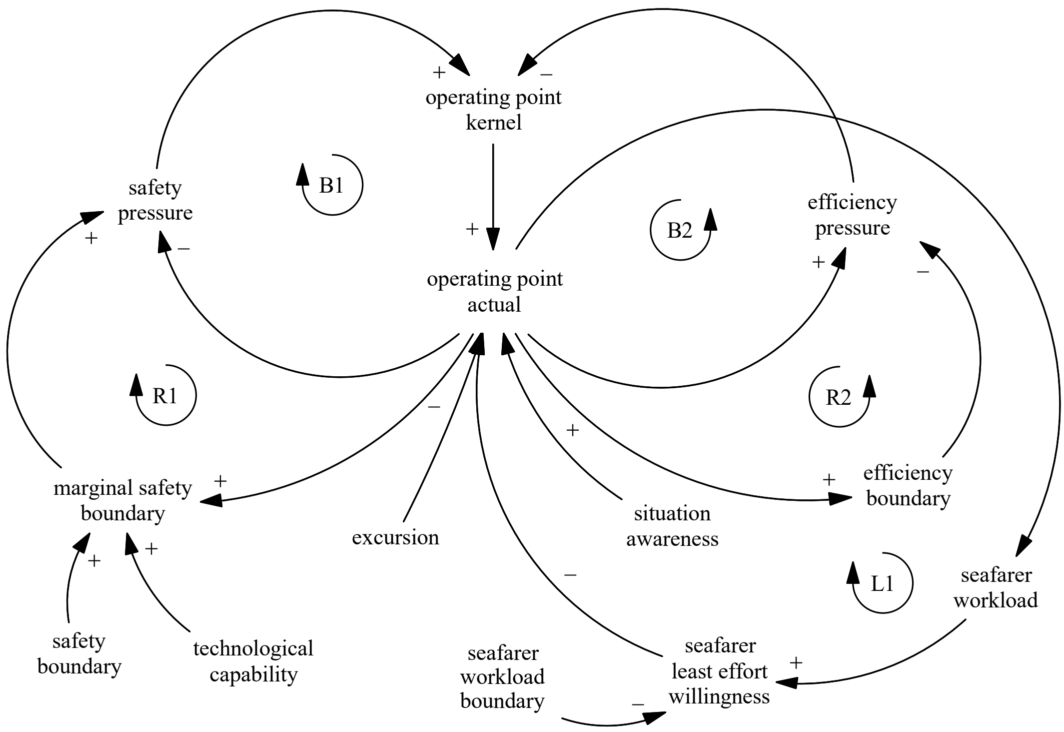


Figure 4.3: Causal loop diagram

as shown in Figure 4.3. Each the arrowhead has + or − symbol at the end where + means the target variable will move in the same direction when the source variable moves, vice versa. In the construction, the first approach is based on the literature study. Second, a safety manager from ocean going shipping company was invited to conduct a hearing about the model. CLD was then updated according to the feedback from the expert.

4.4.1 Operating Point

The operating point in the middle of CLD model demonstrates the operation from the combination of all boundaries. The operating point undergoes continuous movement, driven by a trade-off between effective cost management by the management or onshore side and the worker onboard seafarer who will seek effort gradient. Migration context here refers to the management inclination to seek more suitable positions to gain performance or profit. The analogy of Brownian Movement is used by Rasmussen to explain the phenomenon. The operating point is illustrated as point random that varies within the space of these three boundaries, and the movement will achieve stability when it is minimal and random [64]. However, this movement made the operating point move away from the efficiency and workload boundary, came to closer to the safety boundary; when it cross the safety boundary it leads to an accident.

The operating point is reflected as the result of combining safety pressure and efficiency pressure to determine its magnitude and direction of movement. As previously explained, the + symbol in B1 signifies as the safety pressure increase, the operating point will move to a higher position, indicating greater safety. Conversely, when the safety pressure decreases, the operating point will move to a lower position. The symbol − in loop B2 indicates that as the efficiency pressure increases, the operating point will move to a lower position, implying less safety. From this mechanism, safety and efficiency are operated in similar ways but in different directions. The management strategy of more operational performance is represented as the excursion. The negative arrow input from the excursion suggests that as the excursion value increases, the operating point is moved close to safety boundary.

4.4.2 Safety Pressure

Due to the uncertainty surrounding the location of the safety boundary until an accident occurs, the concept of a marginal safety boundary is introducing a gap from the safety boundary. In the case of ship management companies, it is common to operate near the marginal safety boundary since it offers increased efficiency and reduced workload. When the marginal safety boundary unintentionally crossed without any accident, it is considered as near-miss event, as explained before. In addition, crossing the marginal safety boundary provides the organization with a new perspective that the current operational state is acceptable and thus it is normal to periodically lower the marginal safety boundary as a part of the practice.

Loop R1 in Figure 4.3 represents the safety pressure result from the presence of the marginal safety boundary. When the operating point exceeds the marginal safety boundary, a new lower boundary value is established, positioned lowers to the previous one. This new lower marginal safety boundary reduces the safety pressure as the operating point moves further away from it. With lower operating points, it encourages the operating point to move closer to the safety boundary. Consequently, the feedback loop is the reinforce loop.

4.4.3 Efficiecy Pressure

The company's nature in conducting business is driven by the pursuit of greater profits. Often, for the organization to compromise on safety measures or limits to meet demand from the operational side [65]. The role of manager, such as marine superintendent, involves efficiently allocating resources such as budget and time. However, this allocation process is not without its challenges. Even if the shore staff has extensive experience at sea, the pressure to meet the targets and performance indicators often overlooks the actual onboard situation [69]. These are one of the possible factors in contributing to collecting the pressure to achieve higher efficiency, but unfortunately it also compromise the safety.

According to an extended theory proposed by Morrison and Wears, both marginal safety boundary and efficiency boundary adjusted based on the organization recent experiences and pressure [60]. In the model shown in Figure 4.3, both moveable safety boundary and the efficiency are seen as opposites. In loop R2, the efficiency boundary shifts towards lower levels, indicating increased efficiency but reduced safety due to feedback from operating at lower points without breaching safety or experiencing accident. Conversely, a lower efficiency boundary amplifies the pressure to improve efficiency and further pushes the operating point downwards. Similar to the safety pressure, this loop is categorized as reinforced loop.

4.4.4 Excursion

An excursion results in the operating point moving closer to the safety boundary or less safe condition. Crossing the marginal boundary does not immediately cause an accident, rather it indicates that the current operating point is considered acceptable. Additionally, the organization adjusted the marginal boundary to a lower position in response. Rasmussen explained the factors that influence the direction of the excursion [64]. Additionally, Morrison and Wears explains the factors that influence the magnitude of excursion [60]. The proposed model includes the safety pressure and efficiency pressure as factors that influence the excursion. Increasing safety pressure decreases the excursion and at the same time, increase in efficiency pressure leads to a higher excursion.

4.4.5 Least Effort Workload

In the context of ship operation, where the operator and shore management are geographically and temporally separated, there may be unintended discrepancies in resource allocation. In a broad sense, ship companies prioritize efficiency in staffing by reducing the number seafarers onboard, often opting to eliminate the position of a third mate as a cost-saving measure. This decision places an additional workload on the remaining seafarers [69]. Consequently, maintaining safe operation becomes challenging as the remaining seafarers tend to limit their rest hours and work longer shifts. According to Rajapakse et al report, there have been instances where the work tasks were assigned to seafarers on board without considering their limitation [63]. This results in an allocation that fails to account for the human aspects of seafarers. Increased workloads prompt operators to seek strategies to reduce the burden and achieve a more manageable workload level. These strategies are often seen as shortcuts or workarounds. Furthermore, due to a lack of understanding between shore-based and onboard operators, seafarers frequently withhold specific events or activities from the shore-based staff [69]. Unfortunately, shore management fails to recognize these deviations in the actual work environment.

The crucial point that is emphasized is the presence of additional pressure excursions that occur in the actual field but beyond the estimation of management. This is based on the original model and introduces workload as a distinct factor separate from economics efficiency pressure. As described in loop L1, a movement either increase or decrease in operating point leads to higher workload for front-line operators due to inherent human limitation that define the boundaries of workload. Consequently, a higher workload naturally drives operators to seek strategies that go beyond the workload boundary in order to reduce it and achieve a more suitable workload level. This practice, within the feedback loop, amplifies the significance of the least effort willingness, directly resulting in a decrease in operating point, or in other meaning is less safety.

4.5 Stock and flow diagram construcion

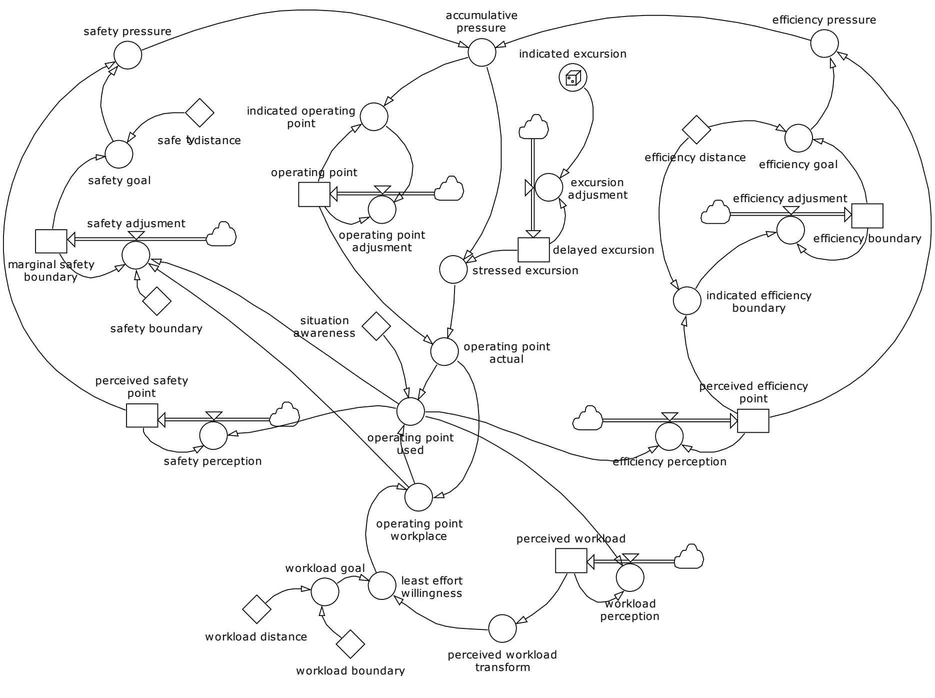


Figure 4.4: Stock and flow diagram

A stock and flow diagram, which was developed based on the CLD, is depicted in Figure 4.4. This diagram visually represents the relationships between different variables, including stocks and flows. The variables in the stock and flow diagram can be broadly classified into two categories: level variables and rate variables. Level variables refer to the quantities that change over the course of the simulation, representing the stocks or accumulations of a particular entity. On the other hand, rate variables are influenced by various factors and subsequently impact the level variables, leading to changes in their values. To facilitate the creation of the stock and flow diagram, as well as the simulation environment, the PowerSim software was employed. This software provided a convenient platform for designing and visualizing the diagram, as well as running simulations to observe the dynamic behavior of the system.

4.6 Simulation result

Since the simulation study in this chapter focuses on construction of theoretical and dynamic pattern, the variables are created most on the assumption. Apart from time, the units used in the simulation are defined as dimensionless (dmnl). The simulation is carried out with a timestep of 1 step, and its definition is independent of real-world time measurements such as days or years. To provide an example, as shown in Figure 4.5, the simulation was executed over a range of 0 to 500 steps, allowing for the observation and analysis of the system's behavior within that timeframe.

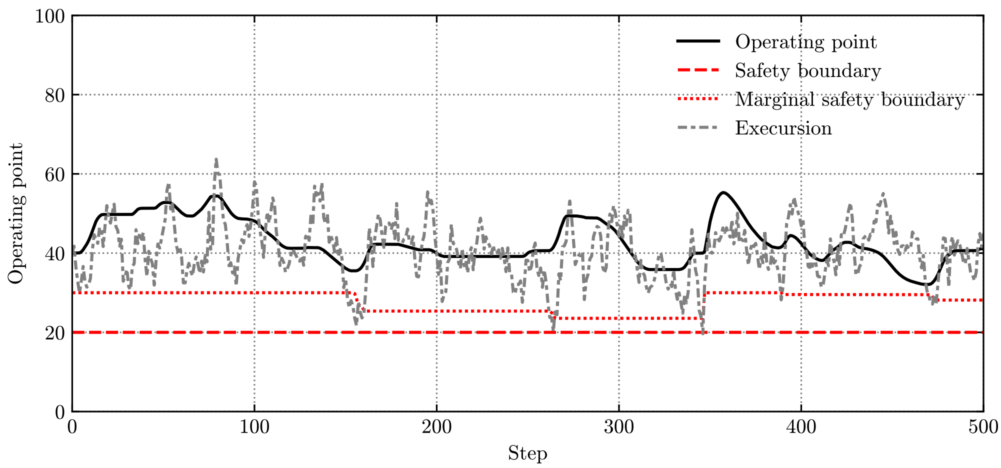


Figure 4.5: Running simulation example

The vertical axis in the graph represents the operation point, where a higher value indicates greater safety, and a lower value indicates less safety. The safety boundary and margin safety boundary are displayed in the lower section of the graph. On the horizontal axis, the operating point is depicted over the time variable, highlighting the dynamic behavior of the system and its changes over time. The objective of this graphical representation is to provide a clear understanding of how the operating point changes and behaves dynamically within the time step dimension.

In the initial of the simulation, the model is run under equilibrium state; the safety and efficiency pressure is equal and the operating poin is on the middle of vertical axis. As shown in Figure 4.5, during the simulation, the excursion that combined with the safety pressure brings the operating point move upward to the higher safety position. In the following step, the efficiency pressure is balancing the condition and moving it to the downward position. Simulating the dynamic properties of marginal safety boundary, the position value of it also moving down as the operating point several times cross the boundary. Last, when the marginal boundary is often brought down and operating point fluctuated in lower part and crossed the safety boundary, accidents happen. In sudden, the safety pressure is increased, and operating point move to the upper part, and the marginal safety boundary is restarted at the initial position.

Situation awareness variable here became the control variable where the simulation is run under various situation awareness level. It demonstrated the condition where the onshore management sufficienct situation awareness of the actual excursion practice onboard operation. For instance, the scenario is divided into the most upper and lower situation awareness as in the Figure 4.6. It can be discriminated, the operating point where the onboard and shore can maintain the situation awreness well, is slightly higher. Instead of created only two scenario, in order to seek its sensitivity, the scenario is divided into five scenarios where the situation awareness is spend accross the 0%, 25%, 50%, 75%, and 100%.

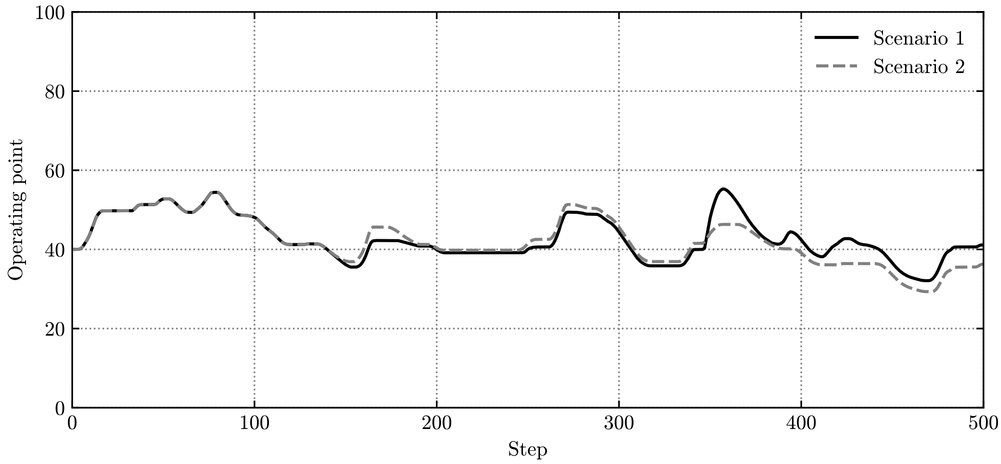


Figure 4.6: Comparison Scenario 1 and Scenario 2

Figure 4.7 shown the simulation result under 100 times running set for each scenario. The vertical axis explains the number of accidents that occur during the duration of simulation. One can notice is that the scenario with higher situation awareness tends to lower its accidents numbers. There is also a pattern where the simulation which has long duration tends to have one way reduction pattern, compare to shorter simulation duration where the there is different pattern with SA in 50% and 75%.

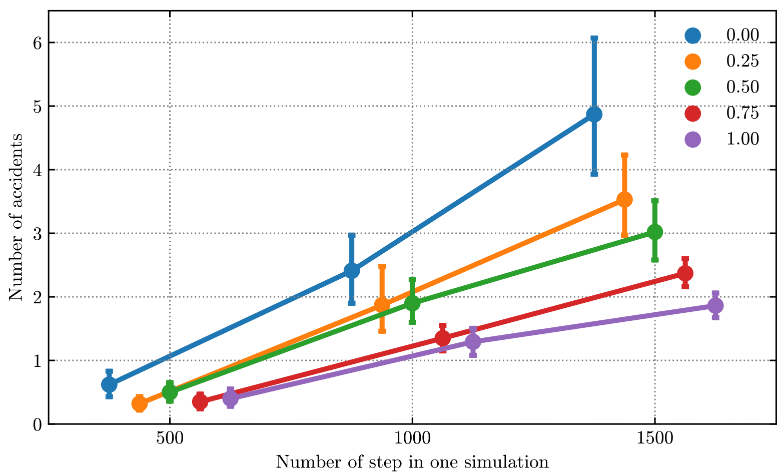


Figure 4.7: Number of accidents

4.7 Discussion

Dynamics safety theory from Rasmussen has been applied in various field of operation, mainly in high-risk organization where the understanding of running operation of organization in the boundary of safety, efficiency and also operator workload is important. In this chapter, the theory is applied in the context of engine resources management between onboard and shore management. Not only qualitative by reproducing and modifying the CLD from the previous study, but also running the quantitative part with creating the stock and flow diagram, running the simulation under differential equation. During the model construction and validating, the expert from safety manager of a shipping company is present to give feedback on the model. Here the validity of the model does not remain only on the theoretical basis of the literature, but also extends into the subject matter expert.

The simulation result shows the ability of the model to follow the Rasmussen theory in most aspects. The trade-off of the safety and efficiency pressure brings the dynamic movement of the operating point. The marginal safety boundary is changed the operating point cross the boundary, and restart when the accident happens. In addition, workload pressure from the operator in this model is reflected on the additional excursion that overshows if the onshore management cannot maintain the situation awareness. In other ways, if the onshore management have this sufficiency understanding about onboard operations, it allows both parties to allocated appropriate safety and efficiency pressure. In the end, the number of accidents can be reduced. Maintaining good coordination between onshore and onboard teams is an essential point in sociotechnical where safety is viewed as control; the leading indicator before accident is known and an intervening action can be conducted.

The system dynamic itself shows the advantage of running the simulation for a long duration where consistency is maintained. However, there are several limitations of using system dynamic to model a safety system. First, as other system dynamic modelling in the sociotechnical system, the numerical assignment is assumed. Therefore, many simplifications on the construction are done in this chapter. More approach study may support the quantitative part, but the dynamic behavior pattern that became interest in system dynamic study may shown the same pattern.

Chapter 5

General Discussion

This chapter summarizes the three studies' approach and result, including the practical and methodological implication of the studies to the area of research. Moreover, the general discussion also mapped the problem during conducting the study. The conclusion is closed by possible future works within the topics.

5.1 Summary and Conclusion

Together with the advancement in ship operation with increasement of automation, also comes a challenge how the operators should adapt to the recent and future operation profile. The situation awareness that introduced as human performance factors, is examined in the scope of human cognition, human reliability, and socio-technical system. The studies are done with the aim of exploring the concept of situation awareness by employing the human factor experiment framework, Bayesian network in human reliability analysis, and dynamic safety theory in system dynamics.

The issue from human-machine interface analysis in engine control highlights the importance of examining situation awareness in the level of human cognition scope. The points from the first study can be summarized below.

- Situation awareness studies in engine resources management is less or limited compare to the navigational operation. Even in the same onboard operation, the work area and characteristic between bridge resources management and engine resources management is different.
- Objective and subjective measurement of situation awareness are applicable in engine resources management with various sensitivity. While subjective measurement is only sensitive to the scenario difficulty or task complexity, objective measurement is also sensitive to the participants experience background.
- The combination objective and subjective with the trade-off of its convenient and customable to the scenario condition, is applicable in evaluation engine resources management training program intended to maintain and improve situation awareness.
- Situation awareness, combined with trust in automation, and workload are beneficial as measurement methods to test the effectiveness of cognition device such as head-worn display. Where the adding addition cognition aid can increase the attention supply, it also increases the attention demand during the utilization.

Considering the situation awareness as the human performance factors is beneficial in constructing the methods to measure the human error analysis. The point from the study in the second chapter is as below.

- Recent human reliability analysis rely on the accident or incident report as the main resources. While the accident report only shows the number of error ratio, not the successful ratio of human conducting the task, the approach using another probability from the simulator experiment is beneficial.

- Proposed structure of Bayesian network consist of performance shaping factor (PSF) as the independent variable from experiment, cognitive function (CF) as the dependent variable, and failure mode (FM) as single variable of error. The structure demonstrates the adaptability of Bayesian network to the experiment design.
- The subject matter expert that previously have big portion in quantify the condition probability distribution is reduced and replace by the experimental result. Thus, the uncertainty can be eliminated.
- The advantage of Bayesian Network is demonstrate with addition step to employ Bayesian interference to update condition probability distribution in the case of the new experimental data exists. It is implied that general PSF database should be useful to combine various resources of simulator data near the future.

In the third study, situation awareness was observed in the scope of sociotechnical system under system dynamics environment. The modelling result of the study brings several conclusions as below.

- The system dynamics model is able to follow the dynamics safety theory by Rasmussen in most aspect, such the operating point of a sosio-technical system is driven by the safety and efficiency pressure.
- The workload aspect that is not modeled in the previous research is included in the model to implicate the workload made by the operator onboard and affected the operating point closer to the safety boundary.
- The role of situation awareness is demonstrate in maintain the undestading of workload level of onboard operation by both parties. Different levels of situation awareness affect the number of accidents produced during the simulation duration.
- Able to allocate appropriate workload level onboard result in reduced number of accidents during the simulation duration. Thus the implementation such strategy to maintain situation awareness between onboard operation and shore management can sustain safe operation.

The limitation remains for each framework used in these three studies. Experiment with simulator must be understand of simplification of real-world environment to able control the independent variables. The simplified scenario also came with the participants who were cadets at the time of experiment are assumed to have enough understanding of the experiment setup. In the second study of Bayesian networks, the limitation remains on the small amount of performance shaping factors that are included in the study. This limitation came across the performance shaping factors that retrieved from the experiment design. The third study limitation came from the system dynamics study that constructed firstly based on the assessment to construct the equation in the system dynamics.

5.2 Future Work

The concept of situation awareness has been examined in the scope of human cognition, human reliability, and socio-technical system. Although each scope has the limitation of study, the future challenges remain. For the scope of human cognition, measurement of situation awareness should be addressed beyond the individual human cognition, namely in team format. Considering this able to cover situation awareness effectiveness in team communication not only in engine resources management, but also between engine and deck department. Moreover, human cognition level in actual engine room situation is the challenge that must be undertaken in the future. In the scope of human reliability, the performance shaping factors should contain any factors than in this study, the failure mode also can be extended to cover other decision making and action in human factors. For socio-technical systems, the challenges remain in changing assumption-based equations and gathering the representative data for safety and efficiency.

The human factors approach that mainly became based on this study is useful to apply in maritime operation towards advancement of automation and autonomous ship era. The changing of seafarer's task with different work environment in the future, from onboard a ship to shore control center as the projection must reflect the limitation of the operator. Therefore, understanding and quantifying human behavior, human reliability such as the situation awareness is such important in dealing with future work environment.

Bibliography

[1] C. D. Wickens and J. G. Hollands, *Engineering Psychology and Human Perforamnce 3rd Ed.* New Jersey: Prentice-Hall, 1999.

[2] N. G. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety*. Cambridge, Massachusetts: The MIT Press, 2011.

[3] K. Wróbel, "Searching for the origins of the myth: 80% human error impact on maritime safety," *Reliability Engineering & System Safety*, vol. 216, p. 107942, Dec. 2021.

[4] M. Grech, T. Horberry, and T. Koester, *Human Factors in Maritime Domain*. CRC Press, 2008.

[5] A. M. Wahl and T. Kongsvik, "Crew resource management training in the maritime industry: A literature review," *WMU Journal of Maritime Affairs*, vol. 17, pp. 377–396, Sept. 2018.

[6] ML. Barnett, CH. Pekcan, and DI. Gatfield, "Content and context: Understanding the complexities of human behaviour in ship operation," in *Proceedings of the Human Factors in Ship Operations Conference. The Royal Institution of Naval Architects*, (London), pp. 1–11, 2005.

[7] R. J. Mumaw, E. M. Roth, K. J. Vicente, and C. M. Burns, "There Is More to Monitoring a Nuclear Power Plant than Meets the Eye," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 42, pp. 36–55, Mar. 2000.

[8] M. Lützhöft, A. Hynnekleiv, J. V. Earthy, and E. S. Petersen, "Human-centred maritime autonomy - An ethnography of the future," in *International Maritime and Port Technology and Development Conference and International Conference on Maritime Autonomous Surface Ships 2019*, vol. 1357, (Trondheim, Norway), Oct. 2019.

[9] Y. Man, M. Lundh, and S. N. MacKinnon, "Managing unruly technologies in the engine control room: From problem patching to an architectural thinking and standardization," *WMU Journal of Maritime Affairs*, vol. 17, pp. 497–519, Dec. 2018.

[10] M. Lützhöft and M. Lundh, "Maritime Application of Control Systems," in *Handbook of Control Room Design and Ergonomics : A Perspective for the Future* (T. Ivergard and B. Hunt, eds.), CRC Press, 2009.

[11] M. H. Lützhöft and S. W. A. Dekker, "On Your Watch: Automation on the Bridge," *Journal of Navigation*, vol. 55, pp. 83–96, Jan. 2002.

[12] L. Maglić and D. Zec, "The Impact of Bridge Alerts on Navigating Officers," *Journal of Navigation*, vol. 73, pp. 421–432, Mar. 2020.

[13] B. Jones, J. Earthy, Ed. Fort, and D. Gould, "Improving the design and management of alarm systems," in *World Maritime Technology Conference*, (2006), p. 7, IMarEST, 2006.

[14] International Society of Automation and American National Standards Institute, "Management of alarm systems for the process industries," Tech. Rep. ANSI/ISA–18.2–2016, ANSI/ISA, 2016.

[15] I. M. Organization, "IMO Resolution A.1021(26) – Code on Alerts and Indicators," Tech. Rep. A 26/Res.1021, International Maritime Organization, 2009.

[16] Y. Cheng, I. Izadi, and T. Chen, "Pattern matching of alarm flood sequences by a modified Smith–Waterman algorithm," *Chemical Engineering Research and Design*, vol. 91, pp. 1085–1094, June 2013.

[17] Danish Maritime Accident Investigation Board, "Marine accident report: Emma Maersk flooding of engine room on 1 February 2013," Tech. Rep. 2013002165, Denmark, 2013.

[18] S. R. Kondaveeti, I. Izadi, S. L. Shah, D. S. Shook, R. Kadali, and T. Chen, "Quantification of alarm chatter based on run length distributions," *Chemical Engineering Research and Design*, vol. 91, pp. 2550–2558, Dec. 2013.

[19] M. R. Endsley, "Situation awareness global assessment technique (SAGAT)," in *Proceedings of the IEEE 1988 National Aerospace and Electronics Conference*, (Dayton, OH, USA), pp. 789–795, 1988.

[20] K. Kuwada, "Tips for effective Engine Management and Maintenance," *P&I Loss Prevention Bulletin*, vol. 49, 2020.

[21] M. R. Endsley, *Designing for Situation Awareness: An Approach to User-Centered Design 2nd Ed.* CRC Press, 2012.

[22] M. R. Endsley, "Measurement of Situation Awareness in Dynamic Systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 37, pp. 65–84, Mar. 1995.

[23] M. R. Endsley, "Situation Awareness Misconceptions and Misunderstandings," *Journal of Cognitive Engineering and Decision Making*, vol. 9, pp. 4–32, Mar. 2015.

[24] M. R. Endsley, "Situation Awareness In Aviation Systems," in *Handbook of Aviation Human Factors* (Garland, Daniel J., Wise, John A., and Hopkin, V. David, eds.), Lawrence Erlbaum Associates Publishers, 1999.

[25] R. A. Grier, J. S. Warm, W. N. Dember, G. Matthews, T. L. Galinsky, J. L. Szalma, and R. Parasuraman, "The Vigilance Decrement Reflects Limitations in Effortful Attention, Not Mindlessness," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 45, pp. 349–359, Sept. 2003.

[26] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "Situation Awareness, Mental Workload, and Trust in Automation: Viable, Empirically Supported Cognitive Engineering Constructs," *Journal of Cognitive Engineering and Decision Making*, vol. 2, pp. 140–160, June 2008.

[27] J. Xu, S. Anders, A. Pruttianan, D. France, N. Lau, J. A. Adams, and M. B. Weinger, "Human performance measures for the evaluation of process control human-system interfaces in high-fidelity simulations," *Applied Ergonomics*, vol. 73, pp. 151–165, Nov. 2018.

[28] C. D. Wickens, S. E. Gordon, and Y. Liu, *An Introduction to Human Factors Engineering*. Longman, 1997.

[29] International Maritime Organization, "IMO Resolution A.947 (23) Human Element Vision, Principles and Goals for the Organization," Tech. Rep. A 23 $\setminus$ Res.947, International Maritime Organization, 2004.

[30] T. Koester, "Human factors and everyday routine in the maritime work domain," in *Human Factors in Transportation, Cmmunication, Health, and the Workplace Human Factors and Ergonomics Society Europe Chapter Annual Meeting*, p. 10, 2001.

[31] M. R. Endsley and D. J. Garland, *Situation Awareness Analysis and Measurement*. CRC Press, 2000.

[32] D. N. Hogg, K. Folles$\oslash$, F. Strand-Volden, and B. Torralba, "Development of a situation awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms," *Ergonomics*, vol. 38, pp. 2394–2413, Nov. 1995.

[33] L. H. Ikuma, C. Harvey, C. F. Taylor, and C. Handal, "A guide for assessing control room operator performance using speed and accuracy, perceived workload, situation awareness, and eye tracking," *Journal of Loss Prevention in the Process Industries*, vol. 32, pp. 454–465, Nov. 2014.

[34] T. Okazaki and C. Nishizaki, "A Study on Situation Awareness of Ship Maneuvering Simulator Training," *International Journal of Emerging Trends in Engineering & Technology*, vol. 03, no. 01, p. 11, 2015.

[35] F. Motz, E. Dalinger, H. Widdel, S. Hockel, and S. MacKinnon, "Comparison of Traditional and Integrated Bridge Design with SAGAT," *International Journal on Marine Navigation and Safety of Sea Transportation*, vol. 3, no. 1, p. 6, 2009.

[36] R. Taylor, "Situational Awareness Rating Technique (SART): The Development of A Tool for Aircrew Systems Design," in *Situational Awareness in Aerospace Opeartions (AGARD-CP-478)*, pp. 3/1–3/17, Neuilly Sur Seine, France: NATO-AGARD, 1990.

[37] S. Selcon, R. Taylor, and E. Koritsas, “Workload or Situational Awareness?: TLX vs. SART for Aerospace Systems Design Evaluation,” *Proceedings of the Human Factors Society Annual Meeting*, vol. 35, pp. 62–66, Sept. 1991.

[38] N. Lau, G. A. Jamieson, and G. Skraaning, “Empirical evaluation of the Process Overview Measure for assessing situation awareness in process plants,” *Ergonomics*, vol. 59, pp. 393–408, Mar. 2016.

[39] Y. Wu, M. Uchida, Y. Sasawaki, M. Kado, G. Lokuketagoda, and T. Miwa, “Changes in work accuracy in a noisy environment : A case of work accuracy in the marine engine room environment,” *Journal of the Korean Society of Marine Engineering*, vol. 41, pp. 461–466, June 2017.

[40] S. Kim, M. A. Nussbaum, and J. L. Gabbard, “Influences of augmented reality head-worn display type and user interface design on performance and usability in simulated warehouse order picking,” *Applied Ergonomics*, vol. 74, pp. 186–193, Jan. 2019.

[41] J.-Y. Jian, A. M. Bisantz, and C. G. Drury, “Foundations for an Empirically Determined Scale of Trust in Automated Systems,” *International Journal of Cognitive Ergonomics*, vol. 4, no. 1, pp. 53–71, 2000.

[42] J. R. Cordon, J. M. Mestre, and J. Walliser, “Human factors in seafaring: The role of situation awareness,” *Safety Science*, vol. 93, pp. 256–265, 2017.

[43] International Maritime Organization, “International convention on standards of training, certification and watchkeeping for seafarers,” tech. rep., International Maritime Organization, London, 2011.

[44] R. Patriarca, “Human reliability analysis: Exploring the intellectual structure of a research field,” *Reliability Engineering and System Safety*, vol. 203, p. 27, 2020.

[45] International Maritime Organization, “Revised Guidelines For Formal Safety Assessment (FSA) For Use in The IMO Rule-making Process,” Tech. Rep. MSC-MEPC.2/Circ.12/Rev.2, 2018.

[46] S. French, T. Bedford, S. J. Pollard, and E. Soane, “Human reliability analysis: A critique and review for managers,” *Safety Science*, vol. 49, pp. 753–763, July 2011.

[47] Z. Yang, S. Bonsall, A. Wall, J. Wang, and M. Usman, “A modified CREAM to human reliability quantification in marine engineering,” *Ocean Engineering*, vol. 58, pp. 293–303, Jan. 2013.

[48] S. I. Ahn, R. E. Kurt, and O. Turan, “The hybrid method combined STPA and SLIM to assess the reliability of the human interaction system to the emergency shutdown system of LNG ship-to-ship bunkering,” *Ocean Engineering*, vol. 265, p. 112643, Dec. 2022.

[49] J. Liu, M. Aydin, E. Akyuz, O. Arslan, E. Uflaz, R. E. Kurt, and O. Turan, “Prediction of human–machine interface (HMI) operational errors for maritime autonomous surface ships (MASS),” *Journal of Marine Science and Technology*, July 2021.

[50] J. Park, W. Jung, and J. Kim, “Inter-relationships between performance shaping factors for human reliability analysis of nuclear power plants,” *Nuclear Engineering and Technology*, vol. 52, pp. 87–100, Jan. 2020.

[51] L. Mkrtchyan, L. Podofillini, and V. Dang, “Bayesian belief networks for human reliability analysis: A review of applications and gaps,” *Reliability Engineering & System Safety*, vol. 139, pp. 1–16, July 2015.

[52] W. Røed, A. Mosleh, J. E. Vinnem, and T. Aven, “On the use of the hybrid causal logic method in offshore risk analysis,” *Reliability Engineering & System Safety*, vol. 94, pp. 445–455, Feb. 2009.

[53] J. Vinnem, R. Bye, B. Gran, T. Kongsvik, O. Nyheim, E. Okstad, J. Seljelid, and J. Vatn, “Risk modelling of maintenance work on major process equipment on offshore petroleum installations,” *Journal of Loss Prevention in the Process Industries*, vol. 25, pp. 274–292, Mar. 2012.

[54] M. Musharraf, D. Bradbury-Squires, F. Khan, B. Veitch, S. MacKinnon, and S. Imtiaz, “A virtual experimental technique for data collection for a Bayesian network approach to human reliability analysis,” *Reliability Engineering & System Safety*, vol. 132, pp. 1–8, Dec. 2014.

[55] M. Aydin, E. Akyuz, O. Turan, and O. Arslan, “Validation of risk analysis for ship collision in narrow waters by using fuzzy Bayesian networks approach,” *Ocean Engineering*, vol. 231, p. 108973, July 2021.

[56] P.-c. Li, L. Zhang, L.-c. Dai, and X.-F. Li, "Study on operator's SA reliability in digital NPPs. Part 3: A quantitative assessment method," *Annals of Nuclear Energy*, vol. 109, pp. 82–91, Nov. 2017.

[57] K. M. Groth, C. L. Smith, and L. P. Swiler, "A Bayesian method for using simulator data to enhance human error probabilities assigned by existing HRA methods," *Reliability Engineering & System Safety*, vol. 128, pp. 32–40, Aug. 2014.

[58] M. Nakamura, T. Miwa, and M. Uchida, "Relationship between characteristics of human factors based on marine accident analysis," in *2015 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, (Singapore, Singapore), pp. 886–891, IEEE, Dec. 2015.

[59] J. Rasmussen, "Human Error and the Problem of Causality in Analysis of Accidents," *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 327, no. 1241, pp. 449–462, 1990.

[60] J. B. Morrison and R. L. Wears, "Modeling Rasmussen's dynamic modeling problem: Drift towards a boundary of safety," *Cognition, Technology & Work*, Apr. 2021.

[61] I. Theotokas, *Management of Shipping Companies*. Routledge, 1 ed., Apr. 2018.

[62] P. M. Barnett, D. Gatfield, and C. Pekcan, "Non-technical skills: The vital ingredient in world maritime technology?," in *Proceedings of the International Conference on World Maritime Technology 2006*, p. 11, 2006.

[63] A. Rajapakse, G. R. Emad, M. Lützhöft, and M. Grech, "A study on time constraints and task deviations at sea leading to accidents – a cultural-historical perspective," *Maritime Policy & Management*, vol. 46, pp. 436–452, May 2019.

[64] J. Rasmussen, "Risk management in a dynamic society: A modelling problem," *Safety Science*, vol. 27, pp. 183–213, Nov. 1997.

[65] J. Reason, *Managing the Risk of Organizational Accidents*. Aldershot: Ashgate, 1998.

[66] N. Leveson, M. Daouk, N. Dulac, and K. Marais, "A Systems Theoretic Approach to Safety Engineering," p. 28, 2003.

[67] R. Cook and J. Rasmussen, ""Going solid": A model of system dynamics and consequences for patient safety," *Quality and Safety in Health Care*, vol. 14, pp. 130–134, Apr. 2005.

[68] J. D. Sterman, *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Boston: Irwin McGraw-Hill, 2000.

[69] H. Sampson, I. Acejo, N. Ellis, L. Tang, and N. Turgo, "The relationships between seafarers and shore-side personnel," tech. rep., Seafarers International Research Centre (SIRC), Cardiff University, 2016.

Appendix A

Simulator Setup

Set A and Set B Freeze Probe Questions

Set	Trouble	Level	Condition	System Parameters in questions
A	M/E L.O Pump Trouble	SA1	Decrease	M/E L.O INLET PRESSURE
		SA1	Decrease	T/C L.O INLET PRESSURE
		SA1	Decrease	M/E L.O INLET TEMPERATURE
		SA1	Decrease	CAMSHAFT L.O INLET PRESSURE
		SA1	False	CENTRAL C.S.W PUMP NO.2
		SA1	Steady	M/E SCAVENGING AIR PRESSURE
		SA2	Steady	M/E AIR COOLER 1 OUTLET TEMPERATURE
		SA2	Steady	M/E SCAVENGING AIR TEMPERATURE
	M/E Jacket CFW Inlet Press	SA1	Decrease	M/E JACKET C.F.W INLET PRESSURE
		SA2	Increase	M/E JACKET C.F.W CYLINDER 1 OUTLET TEMPERATURE
		SA2	Increase	M/E JACKET C.F.W CYLINDER 5 OUTLET TEMPERATURE
		SA2	Decrease	M/E JACKET C.F.W INLET TEMPERATURE
		SA1	Steady	F.O INLET TEMPERATURE
		SA2	Steady	F.O INLET VISCOCITY
		SA1	Steady	M/E EXHAUST GAS CYLINDER 1 OUTLET TEMPERATURE
		SA2	Steady	M/E AIR COOLER 1 OUTLET TEMPERATURE
	Control Air Low Pressure	SA1	Decrease	M/E CONTROL AIR PRESSURE
		SA2	Decrease	CONTROL LINE AIR PRESSURE
		SA2	Decrease	M/E SAFETY AIR PRESSURE
		SA2	Decrease	M/E EXHAUST VALVE SPRING AIR PRESSURE
		SA1	Steady	CENTRAL C.F.W PUMP INLET TEMPERATURE
		SA2	Steady	M/E AIR COOLER 1 INLET TEMPERATURE
		SA1	Steady	CENTRAL C.S.W PUMP OUTLET PRESSURE
		SA2	Steady	M/E AIR COOLER 1 C.F.W OUTLET TEMPERATURE

Set	Trouble	Level	Condition	System Parameters in questions
B	F.O Viscosity High	SA1	Decrease	F.O INLET TEMPERATURE
		SA2	Increase	F.O INLET VISCOCITY
		SA1	Steady	M/E JACKET C.F.W INLET PRESSURE
		SA2	Steady	M/E JACKET C.F.W CYLINDER 5 OUTLET TEMPERATURE
		SA1	Steady	T/C L.O INLET PRESSURE
		SA2	Steady	T/C L.O INLET FLOW
		SA1	Steady	M/E PISTON C.L.O INLET PRESSURE
		SA2	Steady	M/E PISTON C.L.O INLET FLOW
	Air Cooler 1 Contamination	SA1	Decrease	M/E SCAVENGING AIR PRESSURE
		SA1	Increase	M/E EXHAUST GAS CYLINDER 1 OUTLET TEMPERATURE
		SA1	Increase	M/E EXHAUST GAS CYLINDER 2 OUTLET TEMPERATURE
		SA2	Increase	M/E SCAVENGING AIR TEMPERATURE
		SA2	Increase	M/E AIR COOLER 1 OUTLET TEMPERATURE
		SA2	Increase	M/E AIR COOLER 2 OUTLET TEMPERATURE
		SA1	Steady	M/E L.O INLET PRESSURE
		SA2	Steady	M/E L.O INLET TEMPERATURE
	Central CSW Pump Trouble	SA1	Increase	CENTRAL C.F.W PUMP INLET TEMPERATURE
		SA1	Increase	CENTRAL C.S.W PUMP OUTLET PRESSURE
		SA2	Increase	M/E AIR COOLER 1 INLET TEMPERATURE
		SA2	Increase	M/E AIR COOLER 1 C.F.W OUTLET TEMPERATURE
		SA1	Steady	M/E CONTROL AIR PRESSURE
		SA2	Steady	CONTROL LINE AIR PRESSURE
		SA1	Steady	T/C L.O INLET PRESSURE
		SA2	Steady	M/E CAMSHAFT L.O INLET PRESSURE

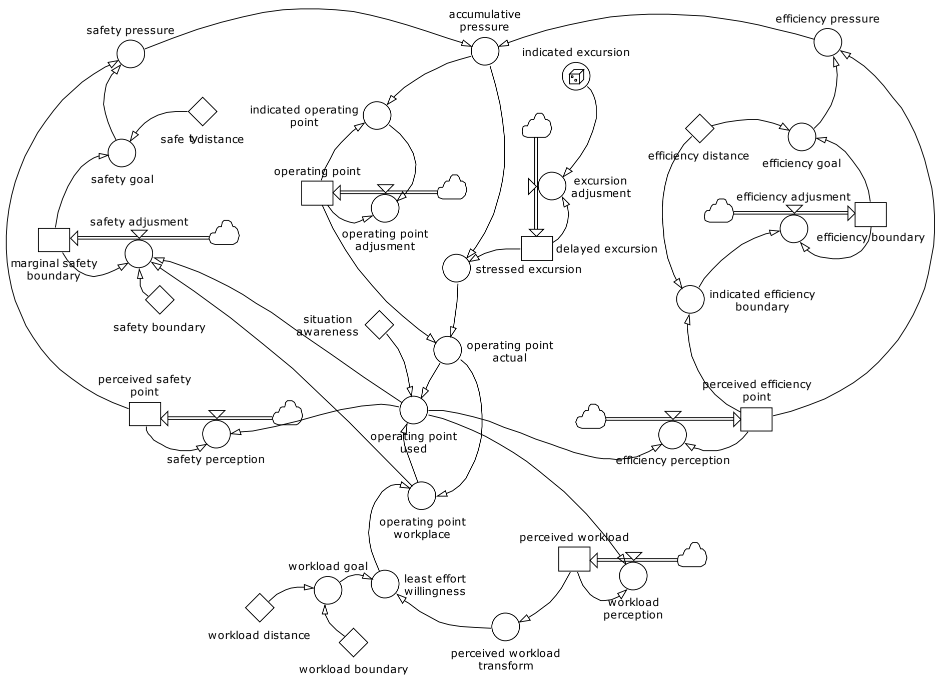
On-going task at the engine room simulator

No	Task	Step
1	Start Transfer Bilge Water, Bilge Well -> Bilge Tank	6
2	Stop Transfer Bilge Water, Bilge Well -> Bilge Tank	6
3	Start Bilge Separator	7
4	Stop Bilge Separator	7
5	Start Transfer Waste Oil, F.O Sludge Tank -> Waste Oil Tank	4
6	Stop Transfer Waste Oil, F.O Sludge Tank -> Waste Oil Tank	4
7	Start Transfer Waste Oil, L.O Sludge Tank -> Waste Oil Tank	4
8	Stop Transfer Waste Oil, L.O Sludge Tank -> Waste Oil Tank	4
9	Start Transfer Waste Oil, Waste Oil Tank -> Waste Oil Service Tank	4
10	Stop Transfer Waste Oil, Waste Oil Tank -> Waste Oil Service Tank	4
11	Start Waste Oil Incinerator	7
12	Stop Waste Oil Incinerator	7
13	Start No.2 F.O Purifier	10
14	Stop No.2 F.O Purifier	10
15	Start No.2 L.O Purifier	10
16	Stop No.2 L.O Purifier	5
17	Start D.O Transfer, D.O Tank -> D.O Settling Tank	5
18	Stop D.O Transfer, D.O Tank -> D.O Settling Tank	5
19	Start F.O Transfer, F.O Tank -> F.O Settling Tank	5
20	Stop F.O Transfer, F.O Tank -> F.O Settling Tank	5
21	Boiler Water Chloride Test	6
22	Makeup Water Chloride Test	6
23	Boiler Water Alkalinity Test	6
24	Condensate Water Ph Test	6
25	Sounding F.W Tank PS	5
26	Sounding F.W Tank SB	5
27	Sounding D.O Tank PS	5
28	Sounding D.O Tank SB	5
29	Sounding F.O Tank PS	4
30	Sounding F.O Tank SB	4
31	Sounding L.O Tank PS	4
32	Sounding L.O Tank SB	4

Appendix B

System Dynamics Source

The appendix explained the simulation environment for the model used in this paper. This includes the equation for all variables in the model, the unit of measure are all dimensionless (dmnl). The setting for all simulation is included. The system dynamic simulation mode is built and runs under Powersim Studio 10, Academic version. The simulation is run using timestep of 1 step.



$$\text{efficiency perception} = (\text{operating point used} - \text{perceived efficiency point})/21$$

The flow of *efficiency perception* is defined by the management as a fraction of difference between the *operating point used* and *perceived efficiency point* with initial setting in 40 dmnl. The fractional adjustment rate is 21 steps.

$$\text{efficiency pressure} = \text{perceived efficiency point}/\text{efficiency goal}$$

$$\text{efficiency goal} = \text{efficiency boundary} - \text{efficiency distance}$$

$$\text{efficiency adjusment} = (\text{indicated efficiency boundary} - \text{efficiency boundary})/21$$

$$\text{indicated efficiency boundary} = \text{perceived efficiency point} + \text{efficiency distance}$$

The value of *efficiency pressure* is defined as ratio between *perceived efficiency point* and *efficiency goal*. Where the *efficiency goal* value is gap set below *efficiency boundary* with *efficiency distance* constant 10 dmnl. *Efficiency boundary* where the initial is set in 60 dmnl, is the product of change by *efficiency adjustment* by fractioned it with *indicated efficiency boundary* in adjustment rate 21 steps. *Indicated efficiency boundary* as a goal is simply 10 dmnl *efficiency distance* from *perceived efficiency point*.

$$\text{safety perception} = (\text{operating point used} - \text{perceived safety point})/7$$

Safety perception is determined as fraction between *operating point used* and the *perceived safety point*. The initial point of *perceived safety point* is set with 40 dmnl, with the fractional adjustment rate is 7 steps.

$$\text{safety pressure} = \text{safety goal} - \text{perceived safety point}$$

$$\text{safety goal} = \text{marginal safety boundary} - \text{safe distance}$$

$$\begin{aligned}\text{safety adjusment} =& \text{IF}(\text{operating point used} < \text{marginal safety boundary},\\ & \text{IF}((\textit{operating point used} < \text{safety boundary})\\ & \text{OR}(\text{operating point workplace} < \text{safety boundary}), (\text{initial marginal safety boundary} -\\ & \text{marginal safety boundary})/1, (\text{operatin point used} - \text{marginal safety boundary})/4, 0)\end{aligned}$$

The value of *safety pressure* is defined as ratio between *perceived safety point* and *safety goal*. The *safety goal* is several distances from the *marginal safety boundary*, in this case is set to 10 dmnl where the initial *marginal safety boundary* is 30 dmnl. *Marginal safety boundary* is a stock variable that will be changed according to *safety adjustment* with fractional adjustment rate is dependent on the state of *operating point used*. If the *operating point used* is below the *m'arginal safety boundary*, the *safety adjustment* becomes the function of fraction between initial *marginal safety boundary* and *marginal safety boundary* with fractional adjustment 1 step. But when the *operating point* used is still above the *safety boundary*, the fractional becomes between it and the *marginal safety boundary* with fractional adjustment 4 step.

$$\begin{aligned}\text{accumulative pressure} =& \text{IF}(\text{safety pressure} > 1, \text{safety pressure}, 1)/\\ & \text{IF}(\text{efficiency pressure} >, \text{efficiency pressure}, 1)\end{aligned}$$

Accumulative pressure acts as the variable to combine and limit the used *safety pressure* and *efficiency pressure*. The value of both pressures is used unless they are above 1 or it means the multiple effect is none.

$$\text{indicated excursion} = \text{NORMAL}(1, 0.2 \times \text{SQRT}((2 - 1/10)/(1/10)))$$

$$\text{excursion adjustment} = (\text{indicated excursion}^{-}\text{delayed excursion})/10$$

$$\text{stressed excursion} = \text{delayed excursion} \times \text{accumulative pressure}$$

Indicated excursion in the Powersim is generated as white noise with function of normal distribution where mean is 1, and standard deviation is set in 0.2. This white noise is then filtered by creating pink noise with *excursion adjustment*, similar to the delay function. *Stressed excursion* acts as excursions that are influenced by *accumulative pressure*.

$$\text{indicated operating point} = \text{operating point} \times \text{accumulative pressure}$$

$$\text{operating point adjusment} = (\text{indicated operating point} - \text{operating point})/5$$

r The *operating point adjustment* has the function to adapt the operating point, with has initial point at 40, to update in approaching *indicated operating point*, with fractional adjustment 5 step. The *indicated operating point adjustment* itself is the multiple effect of *accumulative pressure* to the *operating point*.

$$\text{operating point actual} = \text{operating point} \times \text{stressed excursion}$$

$$\text{operating point workplace} = \text{operating point actual}/\text{MAX}(1, \text{least effort willingness})$$

$$\begin{aligned}\text{operating point used} =& \text{operating point actual} + \text{situation awareness} \times \\ & (\text{operating point workplace–operating point actual})\end{aligned}$$

There are three stages of operating point, first is *operating point actual* where the multiple effect by *stressed excursion* is introduced. *Operating point workplace* where *operating point* is multiple by *least effort willingness* in reduction direction. *Operating point* used is the function of additional difference is added with the variable of situation awareness level. Here the situation awareness has range between 0 and 1.

$$\text{workload perception} = (\text{operating point used–perceived workload})/83$$

$$\begin{aligned}\text{perceived workload transform} =& \text{IF}(\text{perceived workload} \geq \text{initial perceived workload}, \text{perceived workload–} \\ & \text{initial perceived workload}, \text{initial perceived workload–perceived workload})\end{aligned}$$

The flow of *workload perception* is defined by operator as a fraction of difference between the *operating point used* and *perceived workload* with initial setting in 40 dmnl and fractional adjustment is 83 steps. *Perceived workload transform* has function to translate both direction of change on the *operating point* into the increasement of workload.

$$\text{least effort willingness} = \text{perceived workload transform}/\text{workload goal}$$

$$\text{workload goal} = \text{workload boundary–workload distance}$$

The value *least effort willingness* is defined as ratio between *perceived workload transform* and *workload goal*. The workload goal is several distances from the workload boundary, in this case is set to 2 dmnl, where the initial *workload boundary* is 5 dmnl.