Ocean Engineering 251 (2022) 111089

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Application of a SPAR-H based framework to assess human reliability during emergency response drill for man overboard on ships

Sung Il Ahn [a], Rafet Emek Kurt [a,*], Emre Akyuz [b]

[a] Department of Naval Architecture, Ocean and Marine Engineering, University of Strathclyde, 100 Montrose Street, Glasgow, G4 0LZ, UK
[b] Department of Maritime Transportation and Management Engineering, Istanbul Technical University, Tuzla, 34940, Istanbul, Turkey

ARTICLE INFO

Keywords:
SPAR-H
Human reliability assessment
Maritime
Rescue boat drill
Safety
Human factors

ABSTRACT

Emergency preparedness is of paramount importance in successful emergency responses at sea. Therefore, emergency drills are regularly conducted to maintain acceptable levels of emergency preparedness. However, it needs to be considered that emergency drill operations themselves include significant risks, and there is no evidence that these risks are appropriately considered when planning emergency drill operations. Human error is one of the main contributors of accidents during emergency drill procedures. The main question posed is how overall risk, including human errors, during an emergency drill can be correctly evaluated. This paper introduces a new hybrid approach based on the Standardised Plant Analysis Risk Human Reliability Analysis (SPAR-H) method with a fuzzy multiple attributive group decision-making method. The method provides a framework for evaluating specific scenarios associated with human errors and identifies contributors that affect human performance. Estimated human errors are utilised to assess human reliability using a new approach based on a system reliability block diagram. The rescue boat drill procedure for a man overboard is selected to illustrate the method. The findings of this research show each human error probability and its contributing factors per task. As a result, overall reliability of 6.06E-01 was obtained for rescue boat drill operation.

1. Introduction

Research for maritime safety has changed significantly over the past few decades. Understanding these changes can demonstrate how maritime communities can improve maritime safety to reduce maritime industry risk (Luo and Shin, 2019). However, predicting maritime accidents remains a challenge because the cause of accidents consists of various factors, and their contributions remain unclear. Human factors from ship operations are one of the causes of accidents. Studies show that human error is deeply related to maritime accidents (Kristiansen (2013); Ung (2015); Kurt et al. (2016); Akyuz et al. (2018); (Navas De Maya et al., 2019); Antåo and Soares (2019); Ahn and Kurt (2020)) but the terms "human factors" and "human errors" are often used without a clear understanding (Khan, 2008). Human error is directly or indirectly related to several factors called performance shaping factors (PSFs). These PSFs are aspects of human behaviour and contexts that can affect human performance and are often used to derive human error probabilities (HEPs) and identify contributors to human performance. To predict human performance reliability, a contextual description should be provided as discussions of what can happen in a particular situation should be based on a description of the specific circumstances or conditions (Hollnagel (1998); Fujita and Hollnagel (2004)). It is reasonable, then, that human error probability can be affected by a characterisation of the context. Given the unique operating conditions of the shipping industry, in which seafarers face many dangerous situations, proper PSF evaluation is vital in estimating human error.

Emergency preparedness is the most important factor in successful emergency response. Emergency drills in a ship, such as engine room fires, abandoned ships, and oil spills, are regularly conducted in accordance with the Safety of Life at Sea (SOLAS) Convention (IMO, 2018) or their safety management systems. Despite the obviously good intentions of conducting emergency drills, it should be considered whether the risk assessment of these drills, themselves, are being ignored over their benefits. According to the 2001 MAIB report, lifeboats were one of the three significant causes of fatality of seafarers, along with entering confined spaces and falls overboard (Ross, 2006). To prevent accidents occurring during emergency drills, the evaluation procedures for risk of emergency drills should be treated as importantly as the implementation of the emergency drill. Therefore, a customised method for evaluating a ship's condition for the implementation of an emergency drill is needed

* Corresponding author.
E-mail address: rafet.kurt@strath.ac.uk (R.E. Kurt).

https://doi.org/10.1016/j.oceaneng.2022.111089
Received 29 September 2021; Received in revised form 5 February 2022; Accepted 11 March 2022

to avoid accidents caused by emergency drills. This study aims support the ship's decision-making process on the emergency training enforcement through the proposed human reliability assessment framework. Therefore, these outcomes provide valuable contributions to the ship's captain and safety managers to monitor the emergency drills plan and execution.

2. Literature review

Human Reliability Assessment (HRA) is a concern for safety engineers and risk assessment analysts due to fundamental limitations – such as insufficient data –, methodological limitations related to subjectivity of analysts and expert judgment, and limitations caused by uncertainty concerning the actual behaviour of people during accident conditions (Konstandinidou et al., 2006). According to Schröder-Hinrichs et al. (2011), it is more difficult to collect reliable data as human and organisational factors related to accident development and response to emergencies are not reported infrequently. HRAs consists of three important parts: identification, quantification, and reduction (Kirwan, 1994). Many studies have been developed to remedy the existing research gap between the different aspects of HRAs.

Investigators can systematically identify active and latent failures within an organisation that culminated in an accident by reviewing and analysing historical data (Kirwan and Ainsworth, 1992). The Human Factors Analysis and Classification System (HFACS) was developed by Shappell and Wiegmann (2000), and provides a useful tool to assist in the investigation process and to target training and prevention efforts. The HFACS describes human error at the same level as the Reason's Swiss Cheese Model (Reason, 1990) at four levels of failure: organisational influences, unsafe supervision, a precondition for unsafe acts, and unsafe acts. In this model, human error is not the cause of accidents, but is considered a symptom of a bigger problem within the organisation.

Akyuz and Celik (2014) introduced a hybrid model of HFACS combined with cognitive mapping in maritime accident analysis aimed at identifying the distribution of human error. Zhang et al. (2019) proposed a modified model of the HFACS used to analyse ship and icebreaker collisions accident reports and identify the fundamental collision risk factors.

Human error quantification techniques rely either on expert judgment or on a combination of data and psychology-based models, which assess the main impact of human performance (Kirwan, 1994). The major techniques used in the field of human error quantification are discussed below.

The Human Error Assessment and Reduction Technique (HEART) is an HRA technique developed to identify the contribution to human performance and the likelihood of error in a systematic and repeatable way (Bell and Williams, 2016). This method based on the general principle that every task has a nominal failure probability. Influencing each of these tasks are varying levels of Error Producing Conditions (EPCs), that can affect human performance in systems operations (Uflaz et al., 2022). The method provides users with human reliability data that can be modified to be specific to their risks. HEART is a relatively quick and straightforward method that is applicable to any industry where human reliability is important. Despite the occasional opposition, it is not a limited method to the nuclear field. Noroozi et al. (2014) applied HEART analysis to human error during maritime maintenance operations. Akyuz and Celik (2016) also introduced the application of HEART combining the Analytic Hierarchy Process (AHP) to the case of a cargo loading operation in an oil/chemical tanker ship for human error probabilities estimation. Islam et al. (2017a) developed an operational-specific methodology based on HEART used to capture unique features of maritime environments and operations and applied it to the maintenance procedures of a maritime engine exhaust turbocharger and a condensate pump on offshore oil and gas facilities.

The Cognitive Reliability and Error Analysis Method (CREAM) was first developed by Hollnagel (1998), and was used to predict human performance reliability. The human error probability can be determined directly from a characterisation of the context, based on a description of the specific circumstances or conditions (Fujita and Hollnagel (2004)). Since the introduction of CREAM, numerous follow-up studies have been conducted by researchers from different disciplines to provide a greatly advanced CREAM method. Ung and Shen (2011) proposed a systematic procedure to compute probabilities of operator action failure in and Lee et al. (2011) suggested a customised CPC called Cognitive Speaking Process, which focuses on communication error in a nuclear plant. Meanwhile, a simplified CREAM method was introduced to provide an easily practicable process to numeric output results, which can be applied to both the basic method and extended method. He et al. (2008), Akyuz (2015), Akyuz and Celik (2015), and Xi et al. (2017) utilised a simplified CREAM method which was developed to provide an easily practicable process to numeric output results. However, numerous assumptions were made to estimate these numerical results, which may have introduced uncertainty. For example, it assumed that if different scenarios have an equal difference of negative and positive impacts, then they will have the exact failure probabilities.

The Success Likelihood Index Method (SLIM) is a technique used in the HRA field to evaluate the probability of a human error throughout the completion of a specific task. SLIM provides a set of models for the factors that influence human error during commonly occurring activities, including alarm response, actions, checking, information retrieval, and communication. SLIM is a decision-analytic approach to quantifying PSFs using expert judgement, and factors related to an individual, environment, or task are likely to have either a positive or negative impact on human performance. These factors are used to derive a Success Likelihood Index (SLI), a form of preference index that is corrected for existing data to derive a final Human Error Probability (HEP). The PSFs which must vitally be considered are chosen by experts and are namely those factors that are regarded as most significant concerning the context in question. Abbassi et al. (2015) presented combined HRA by integrating the SLIM with the Technique of Human Error Rate Prediction (THERP) for an offshore condensate pump maintenance task. This method is based on different data collections to leverage the quantitative data provided by THERP and generate human error data in SLIM if the data is not available. Islam et al. (2017b) developed a monograph to assess the likelihood of human error in maritime operations that is applicable to the instant decision making by using SLIM to estimate HEP.

The Standardised Plant Analysis Risk Human Reliability Analysis (SPAR-H) method was developed to estimate the human error probabilities associated with operator and crew actions and decisions in response to initiating events at commercial U.S. nuclear power plants by Blackman et al. (2008). The SPAR-H method has been applied to human error-related research in various industries. As an example, Jahangiri et al. (2016) applied the SPAR-H method to analyse and quantify the potential human errors and extract the required measures for reducing the error probabilities in the permit to work system in a chemical plant. In the petroleum industry, the Petro-HRA method, which used the SPAR-H method as the basis for the quantification model, has been developed to analyse human actions as barriers in major accidents and applicability of human reliability analysis methods (Bye et al., 2017). The Petro-HRA project provides guidance on the comprehensive process of HRA as well as human error quantification. In the Maritime industry, Parhizkar et al. (2021) applied SPAR-H method to estimate the effect of performance shaping factors on the human error probability for the probabilistic risk assessment of Decision-making in Emergencies of the dynamic positioning drilling unit.

The Fuzzy theory and Bayesian network are commonly integrated into HRA techniques to develop more advanced human reliability methods. Fuzzy logic is utilised to convert qualitative data to quantitative data and opinion aggregation from multiple groups with combined HRA techniques, while Bayesian networks are applied to consider dependency and weighting among PSFs. These techniques contribute to

enhance the consistency of research and minimise subjectivity and ambiguity during expert judgment and provide instant calculations of human error probability. Examples of past studies predicting human error using these hybrid methods are as follows. Konstandinidou et al. (2006) developed a fuzzy modelling system to estimate the probability based on CREAM methodology. This developed fuzzy logic consisted of 9 input variables from Common Performance Condition (CPC) and applied an if-then knowledge-based fuzzy inference system to predict an accurate number of human errors from qualitative opinions. Yang et al. (2013) proposed a modified CREAM by incorporating fuzzy evidential reasoning and a Bayesian network based on fuzzy inference logic for an oil tanker's Cargo Oil Pumps shutdown scenario. This approach consisted of a multiple-input multiple-output rule concept which was sensitive to the minor changes of fuzzy input. Ung (2015) considered the weight of each CPC in applying a fuzzy CREAM method to the maritime scenario to estimate human error probabilities. Zhou et al. (2017a) developed the eight customised CPCs to reflect better the essential aspects of the contextual condition for maritime operations, and applied the weighting of the CPCs by employing the Fuzzy Analytical Hierarchy Process (FAHP). Akyuz (2016) introduced another HRA technique application, utilising SLIM with the fuzzy sets during the weighting process of PSFs, to abandon-ship procedures. Erdem and Akyuz (2021) assessed the potential contribution of human errors in the maritime industry by using SLIM with interval type-2 fuzzy sets, which effectively addressed subjectivity in the process of using experts' judgments for loading operations in a container ship. In SPAR-H studies, Groth and Swiler (2013) and Groth et al. (2014) introduced the BN version of SPAR-H. Mirzaei Aliabadi et al. (2019) utilised the Bayesian network (BN) to improve the probability of human error allocated by the existing HRA method to identify and analyse human errors in the pegging work of the gas transmission plant.

Once human reliability is assessed, the next step is a human error representation which utilises modelling to carry out a risk assessment for a reduction of error. Some studies have demonstrated a risk assessment combining the human reliability assessment methods. For example, Zhou et al. (2017b) utilised the CREAM method with a modified fault tree model for LNG spill accidents during LNG carriers' handling operations for risk assessment. Ung (2019) applied Fault-tree analysis where a modified Fuzzy Bayesian network-based CREAM was applied to a risk assessment of human error contribution in oil tanker collisions.

Through the above literature review, cases of application of HRA research in various industries, including the maritime, were examined. Although various HRA techniques are applied to the maritime cases to enhance safety, the following research gaps are identified for HRA application in the Maritime industry. Firstly, these HRA techniques mainly focus on quantifying human errors while they do not deal well with the dependency among tasks. Furthermore, the previous HRAs do not address how to incorporate each failure event into the system structure in a detailed method. Therefore, systematic modellings, including human error, need to be developed. Secondly, the issue of uncertainty and inconsistency in expert judgment arising from the process of quantifying human errors remains to be improved despite efforts in previous studies. Finally, as mentioned above, SPAR-H is a technology developed for the nuclear industry, so the provided PSFs are from the operation of the nuclear industry and need to be customised for application to specific operations at maritime.

In this regard, the purpose of this study is not only to create a more accurate framework for quantifying human errors with contributing factors, but also to provide a new approach to human reliability modelling. To achieve the research objective, modified SPAR-H is employed to estimate human error probability. The fuzzy opinion aggregation method contributes to enhancing the consistency of research and minimising subjectivity and ambiguity. Furthermore, the Reliability Block Diagram (RBD) analysis is used to model human reliability to consider system configuration and dependency among tasks. The procedures of a rescue boat drill for a man overboard in a specific context defined by a scenario is selected to present human error probabilities for each task and human reliability for the entire procedures.

3. Methodology

This study presents a modified SPAR-H technique with a fuzzy opinion aggregation to quantify human error probability in the event of maritime emergency procedures. Before introducing the proposed new method, the following section gives a brief introduction to SPAR-H.

3.1. SPAR-H

The Standardised Plant Analysis Risk Human Reliability Analysis (SPAR-H) method was developed to estimate the human error probabilities associated with operator and crew actions and decisions in response to initiating events at commercial U.S. nuclear power plants by Blackman et al. (2008). In the SPAR-H approach, calculation of HEP rates is especially straightforward, starting with pre-defined nominal error rates for cognitive versus action-oriented tasks, and incorporating performance shaping factor multipliers upon those nominal error rates (Blackman et al., 2008). In SPAR-H, the process of human error quantification is summarised as follows: The SPAR-H categorises human failure events (HFE) as either Diagnosis tasks or Action tasks or combined Diagnosis and Action (Whaley et al., 2011). Once HFEs are categorised, analysts identify factors that affect human performance in both a positive and negative way to support qualitative evaluation. This process can be supported by reviewing SPAR-H performance shaping factors. These factors include eight PSFs: time available, stressors, experience and training, complexity, ergonomics including human-mechanical interfaces, procedures, fitness for duty, and work processes. When the PSFs level is specified, the final HEP is the product of the nominal HEP and the composite multipliers of PSF, with the following equations:

$$\text{HEP} = \text{Diagnosis Error} + \text{Execution Error} \tag{1}$$

$$\text{Diagnosis Error} = \text{Nominal Diagnosis Error x Composite Multipliers of PSFs} \tag{2}$$

$$\text{Execution Error} = \text{Nominal Execution Error x Composite Multipliers of PSFs} \tag{3}$$

When more than three negative PSFs, human error probability needs to be adjusted by equation (4).

$$Adjusted\ \ HEP = \frac{0.01^{*}\prod multipliers\ of\ PSFs}{0.01^{*}(\prod multipliers\ of\ PSFs - 1) + 1} + \frac{0.001^{*}\prod multipliers\ of\ PSFs}{0.001^{*}(\prod multipliers\ of\ PSFs - 1) + 1} \tag{4}$$

3.2. Proposed approach

This section proposes a hybrid approach that combines SPAR-H with the fuzzy theory to assess human reliability in maritime onboard procedures. To minimise the subjectivity and variability of experts, we

adapt and customise fuzzy multi-attribute group decision-making methodologies by Ölçer and Odabaşi (2005) for an opinion aggregation. The context of critical maritime scenarios may include factors like the human-machine interfaces, the complexity of the task, working conditions, and crew training levels. However, different operations are not carried out in the same environment. Therefore, the characteristics of each task and the factors that affect its performance should be evaluated individually. For this reason, the Petro-HRA method by Bye et al. (2017) – based on SPAR-H by Blackman et al. (2008) – is selected as an appropriate framework for evaluating ship offshore emergency procedures. This is because the SPAR-H method helps measure the effectiveness of performance shaping factors on human performance for individual tasks to estimate human errors, while the Petro-HRA provides a comprehensive quantitative risk assessment framework for whole procedures. The flowchart of the proposed approach is shown in Fig. 1.

3.3. Scenario definition

The scenario defines the scope and boundaries of the analysis and is used as the underlying data for subsequent qualitative and quantitative analysis (Bye et al., 2017). This step focuses on describing the context of individual tasks throughout the whole process. The main objective of scenario development is to create a more detailed description of the event sequence to identify potential human errors better and understand the operational context. The scenario in this paper includes detailed information such as the tasks performed, individuals responsible and their roles, the task location (indicating the working conditions and the external environmental conditions), and the equipment used with their interfaces.

3.4. Task analysis

The goal of task analysis in this research can be defined as simply subdividing the functions into tasks, tasks into subtasks, and subtasks into human actions. A task analysis describes the steps performed as part of the activity, providing a method of systematically organising the information collected about the task (Sezer et al., 2022; Bye et al., 2017). In this paper, two different task analysis methods are utilised. First, a hierarchical task analysis is performed to define the task on the procedure's primary goal, along with subtasks to address the specified duties that the operator should complete. Second, the HTA provides a graphical overview of the tasks involved in the analysis scenario. However, hierarchical task analysis is not sufficient to provide appropriate information in the context associated with the tasks. Therefore, a Tabular task analysis is adopted to provide more information for experts' judgment and better organise data.

3.5. Deriving and rating PSF

This section begins with the definition of the PSF and describes the step-by-step process of implementing expert evaluations and representing the consensual results in the corresponding PSF multipliers.

3.5.1. Define PSFs and guidance for PSF ratings

The selection of PSFs that affect human performance and their assessment criteria should change depending on context. Therefore, the PSFs definition, levels provided by Whaley et al. (2011), were refined by maritime experts with customised guidance used to establish and rate characteristics in onboard rescue drills. The provided description of PSFs

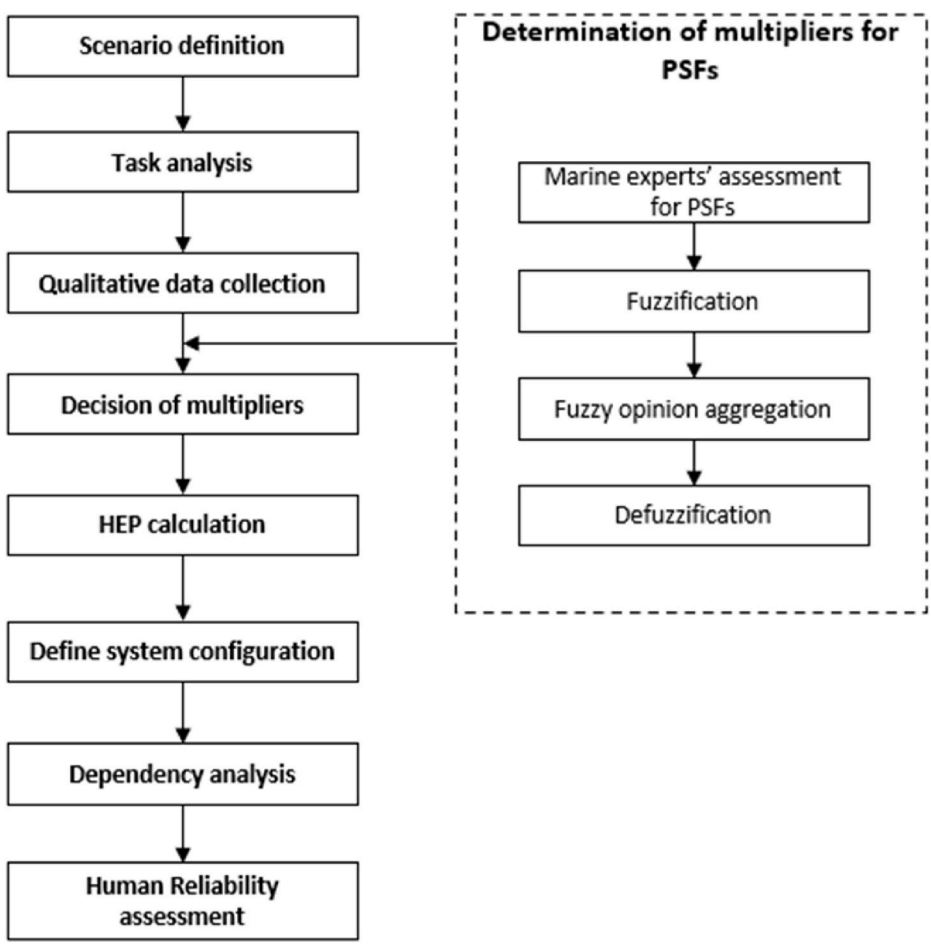

Fig. 1. Flow chart of the proposed approach.

should be as clear as possible to experts to determine the appropriate PSFs rating for the task being analysed, while preventing them from mechanically selecting the PSF rating. The criteria for PSFs are set in this section through expert consensus before judgment for PSFs, considering these two opposing aspects at the same time.

i) Human-machine interface (PSF_1)

The Human-Machine Interface (HMI) PSF refers to the quality of equipment, controls, hardware, software, monitor layout, and the physical workstation layout, where the operator/crew receives information and carries out tasks (Bye et al., 2017). Human-machine interfaces should be properly evaluated for the tasks required from two perspectives: interfaces for diagnosis such as monitors and visible & audible alarms, and interfaces for execution such as switch buttons, levers, and keyboard(s).

ii) Threat Stress (PSF_2)

Threat stress refers to the expectation or fear of physical or mental harm (Salas et al., 1996). Examples of situations that can cause threat stress on a ship include fear felt by the worker in a confined space, or fear that a lifeboat could fall when it is suspended high above the surface of a davit fall.

iii) Level of experience or training skill (PSF_3)

This PSF refers to the experience and training of the operator(s) involved in the task and should focus on satisfying the experience/skill required by the assigned task, which is identified through task analysis rather than measuring the skills of the worker(s) in a wide range of areas.

iv) Procedures (PSF_4)

The procedures PSF represents the existence and use of formal operational procedures for the task and includes user manuals and instructions for machine and software operations for the task. The procedure is assessed from the following perspectives: whether all required procedures are in place; whether the procedures are easily accessible and visible from the workplace; whether the procedure/manual/instruction contains enough content to perform the task; and whether the content is unambiguous and is easy to understand linguistically and graphically.

v) The complexity of the task (PSF_5)

Task Complexity refers to how difficult the task is to perform in the given context. The degree of complexity is measured using different information, including physical and mental hardness, the number of goals, and the number of steps.

vi) The working condition (PSF_6)

The working condition refers to the physical variables in which the work is performed (e.g., temperature, humidity, vibration, noise level, allowable space, and intensity of light). This affects mental state by causing certain moods and emotions. For example, suppose the workplace moves even though it occupies the same space (e.g., a lifeboat). In that case, the different characteristics of lifeboats at stowed positions and lifeboats on the sea should be reflected.

vii) Environmental condition (PSF_7)

Environmental Conditions refer to the state of the ship's environment, including weather conditions, sea conditions, and the time of day. The effects of environmental conditions should be evaluated differently depending on the location (e.g., control room, boat on the sea, etc.).

viii) Time pressure (PSF_8)

Time pressure indicates the amount of pressure that requires the operation to be performed in time. Thus, the time pressure depends on the available time compared to the minimum time taken. Examples of negative effects include scenarios where: the event to avoid has already occurred or where it is too late to recover within a specified period; a slight delay in time has serious negative consequences; the operator must complete the operation before starting the next sequence operation; and where the task must be performed simultaneously with other tasks or at a specific time.

ix) Ship safety management system (SMS) and supports (PSF_9)

The PSF refers to Safety, Work, and Management Support consists of three related factors: 1) adequacy of established SMS; 2) the degree of implementation of SMS; and 3) the degree of support offered by the company to perform tasks.

3.5.2. Adaption of fuzzy theory

The selected PSFs have linguistic variables that negatively or positively represent the level of PSFs dealing with expected impacts on performance reliability. In conventional SPAR-H, monolingual variables are determined with 100% faith in the relevant PSF evaluation. However, a limited number of language scales are insufficient to reflect the impact of PSF on human confidence in real-world situations (Ahn and Kurt, 2020). To better describe the impact of PSFs, fuzzy sets are employed because they are useful practices in dealing with the ambiguity of human error detection problems (Akyuz, 2016). Each PSF connects eight triangular fuzzy sets to illustrate the impact of each PSF as shown in Fig. 2.

The triangular fuzzy set expressed as (a, b, c) and membership function μ(x) for a linguistic variable is obtained as follows.

$$\mu(X) = \begin{cases} \frac{x-a}{b-a}, & a \le x < b \\ 1, & x = b \\ \frac{c-x}{c-b}, & b < x \le c \\ 0, & \text{Otherwise} \end{cases} \quad \text{where } a \le b \le c \tag{5}$$

3.5.3. Experts' judgment and fuzzy opinion aggregation

A group of experts are asked to evaluate the level of each PSF considering the characteristics of each task. A linguistic scale for PSF levels and their corresponding fuzzy set are developed and provided in Table 1.

The purpose of applying the fuzzy opinion aggregation in Fig. 1 is to translate the experts' multiple qualitative assessments of PSF ratings into a single aggregated opinion with fuzzy opinion and convert it into a crisp value through defuzzification. The modified opinion aggregation procedure, adapted from (Ahn and Kurt, 2020), is made based on a fuzzy multiple attributive group decision making methodology by Ölçer and Odabaşi (2005) as follows:

(a) Calculating the degree of agreement (Similarity)

Assume that the fuzzy set selected by experts A and B as A= (a_1, a_2, a_3), B= (b_1, b_2, b_3), and A and B are standardised fuzzy sets. Here, S (A, B), which is the degree of similarity between A and B, is measured by equation (6):

$$S(A,B) = 1 - \frac{|a_1 - b_1| + |a_2 - b2_1| + |a_3 - b_3|}{3x\ \ Nmax} \tag{6}$$

N_{max} is seven because a maximum level of PSF is defined as seven.

(b) Calculating the average degree of agreement (AA)

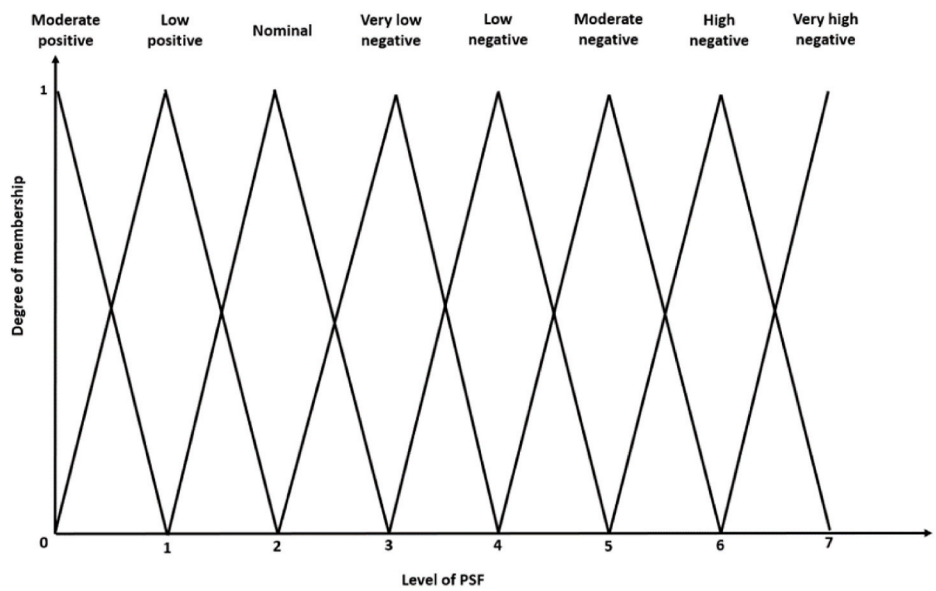

Fig. 2. Fuzzy membership for PSFs.

Table 1
Evaluation of PSFs influence on performance with fuzzy sets.

Level of PSF	Fuzzy set	Number of PSF(N)	SPAR-H multiplier	Interpolated Multiplier
Extremely high negative[a]	–	–	HEP = 1	HEP = 1
Very high negative	(6, 7, 7)	N = 7	50	50
		6 < N < 7	–	25N-125
High negative	(5, 6, 7)	N = 6	25	25
		5 < N < 6	–	15N-65
Moderate negative	(4, 5, 6)	N = 5	10	10
		4 < N < 5	–	5N-15
Low negative	(3, 4, 5)	N = 4	5	5
		3 < N < 4	–	3N-7
Very low negative	(2, 3, 4)	N = 3	2	2
		2 < N < 3	–	N-1
Nominal/not applicable	(1, 2, 3)	N = 2	1	1
		1 < N < 2	–	0.5N
Low positive	(0, 1, 2)	N = 1	0.5	0.5
		0 < N < 1	–	0.4N + 0.1
Moderate positive	(0, 0, 1)	N = 0	0.1	0.1

[a] If one (or more) PSFs are an extremely high negative case, then the HEP for the corresponding task shall be set to 1 regardless of any other multipliers for the other PSFs.

Let's define $AA(Ex_i)$ as the i_{th} average degree of agreement between $expert_i$ and $expert_j$, and this can be calculated by equation (7):

$$AA(Ex_i) = \frac{1}{D-1} \sum_{\substack{i=1 \\ i \neq j}}^{D} S(Ex_i, Ex_j) \tag{7}$$

Where D is the number of experts.

(c) Calculating the relative degree of agreement (RA)

Let's define $RA(Ex_i)$ as the i-th relative degree of agreement which can be calculated by equation (8):

$$RA(Ex_i) = \frac{AA(Ex_i)}{\sum_{i=1}^{D} AA(Ex_i)} \tag{8}$$

(d) Calculate the consensus degree coefficient (CC)

Let us define $CC(Ex_i)$ as the consensus degree coefficient for i-th expert which can be calculated by equation (9):

$$CC(Ex_i) = \beta * w_i + (1-\beta) * RA(Ex_i) \tag{9}$$

Where β is a relaxation factor between 0 and 1. Note that a Homogeneous group of experts can be calculated by assigning β as 0.

(e) Calculating the aggregation result of the fuzzy opinion (R_{AG})

The aggregated fuzzy set R_{AG} can be calculated using the following equation:

$$R_{AG} = \sum_{i=1}^{D} CC(Ex_i) * P(Ex_i) = (S_1, S_2, S_3) \tag{10}$$

(f) Defuzzification.

Finally, an aggregated fuzzy set R_{AG} for each PSF is converted to a crisp value by a centre of gravity (COG) method, demonstrated below:.

$$\text{Defuzzified rating of PSF} = \frac{S1 + S2 + S3}{3} \tag{11}$$

3.5.4. Calculation of multipliers from the defuzzied PSFs ratings

The SPAR-H output in Table 1 provides eight multipliers for eight different PSF ratings, but information on multipliers between integer intervals is not available. To make the multiplier a continuous number, it is assumed that the function follows a linear pattern between adjacent PSF ratings, as shown in Fig. 3. The seven functions of linear lines between points can be calculated based on the PSF rating and their multiplier, respectively and this idea using linear lines adopted from research for dynamic probabilistic risk assessment of decision-making for dynamic positioning drilling unit (Parhizkar et al., 2021). For example, a value corresponding to the multiplier for the rating of numbers N between 6 and 7 can be interpolated by using Equation (25) N-125. Seven linear functions for each interval are listed in Table 1. This increases sensitivity by providing a corresponding multiplier for consecutive numbers obtained using fuzzy, rather than integers, resulting in an accurate value.

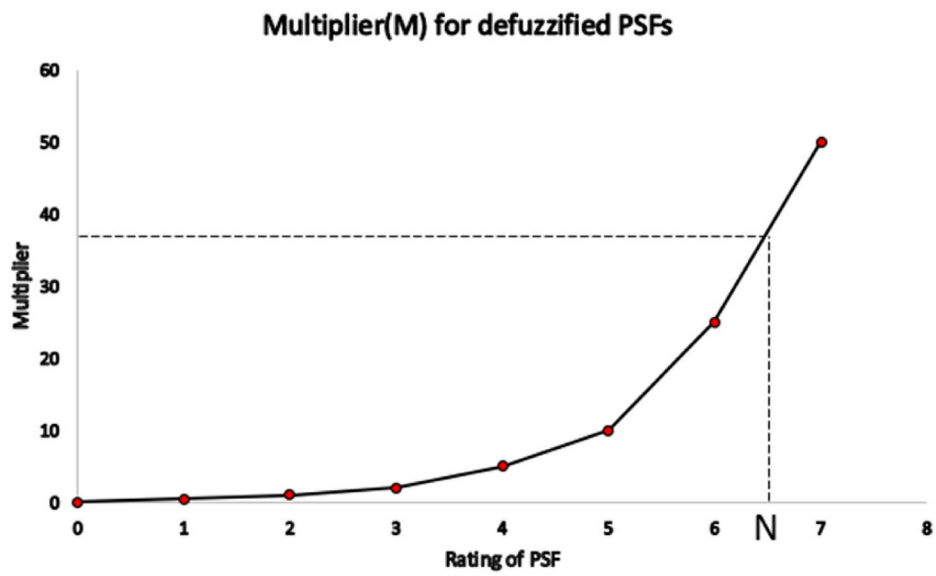

Fig. 3. Multiplier for defuzzified PSFs level.

3.6. Human error quantification

According to SPAR-H, human error probability is calculated as the sum of diagnosis error and execution error. Each error has a nominal failure probability of 0.01 and 0.001, respectively. However, human error probability in this paper is calculated using Equation (12), as the 'OR' gate calculation is more reasonable and consistent for computing other failure events than the sum of diagnosis and execution errors.

$$\text{HEP} = 1 - (1 - \text{Diagnosis Error}) \times (1 - \text{Execution Error}) \quad (12)$$

3.7 Modelling of human reliability assessment.

Once the human error probability of a sub-task is derived, several rules need to be formulated to calculate the total probability of failure of the entire task. The equations listed in Table 2 are used to obtain the total human reliability in the entire procedure. For sub-task with low or no dependency, failure probability is derived from the multiplication of HEP for sub-task$_i$ in parallel systems, and failure probability is derived from the sum of HEP for sub-task$_i$ in series systems. For parallel subtasks with high or complete dependence, the minimum HEP of all sub-works is used. That is, since the task succeeds when any of the sub-tasks is successful, the probability of success of the entire task is assigned to the highest probability of success of the sub-task. For sequential sub-tasks with high or complete dependence, the maximum HEP of all sub-works is used. That is, when one of the sub-tasks fails, the highest probability of failure of the sub-task is assigned as the probability of failure of the entire task. The mentioned method proposed by (He et al., 2008) provides a simple and effective way to calculate human reliability, but the following assumptions should be applied. First, all sub-tasks constituting the task should be connected to either parallel or serial systems. Second, the level of dependence on all sub-tasks should be the same within the task. However, since a combination system of series and parallel cannot be applied, there is a limit to its application to actual cases. In addition, the level of dependence of sub-tasks cannot always be assumed to be the same for all cases within the task and may be different. Therefore, this paper proposes a new approach using a Reliability Block Diagram (RBD) by assuming each task and sub-task are system components for this HRA modelling. Details will be explained in conjunction with the cases study illustrated in Section 4.5.

Table 2
Calculating the Human error probability from HEPs of its sub-tasks (He et al., 2008).

System description	System sub-task dependency	Notation for task HEP & Reliability
Parallel system	High dependency	$HEP_{Task} = Min\{HEP_{Sub-task\ i}\}$ or (13) $R_{Task} = Max\{R_{Sub-task\ i}\}$ (14)
	Low or no dependency	$HEP_{Task} = \prod(HEP_{Sub-task\ i})$ or (15) $R_{Task} = 1 - \prod(1 - R_{Sub-task\ i})$ (16)
Series system	High dependency	$HEP_{Task} = Max\{HEP_{Sub-task\ i}\}$ or (17) $R_{Task} = Min\{R_{Sub-task\ i}\}$ (18)
	Low or no dependency	$HEP_{Task} = 1 - \prod(1 - HEP_{Sub-task\ i})$ (19) $\approx \sum(HEP_{Sub-task\ i})$ or $R_{Task} = \prod(R_{Sub-task\ i})$ (20)

4. Case study

For an illustration of the proposed approach, both scenario and procedures for the man overboard rescue drill during ship navigation have been selected. This scenario was developed based on the real shipboard rescue drill observed. The scenario of a man overboard rescue drill is described in section 4.1 to assess PSFs and estimate human error probabilities. The hierarchical task analysis and tabular task analysis are conducted respectively and are described in section 4.2.

4.1. Scenario definition

The scenario for a rescue boat drill for a man overboard on a ship is described as an illustration of the proposed method. One applied assumption is that a watchkeeping man is able to observe a man overboard. According to SOLAS regulation III/14.1 (IMO, 2018), the rescue boat should be launched in no more than 5 min.

The ship was under navigation in the open sea from the port of Nagoya to the port of Tokyo for cargo loading. The temperature was 32 °C, and the humidity was 70%. The wind speed was moderate, and the current speed was relatively high. The vessel was a newly constructed container ship, G/T 6,500 TEU, and the overall vessel condition was described as good. The ship's management company has managed a total of 130 vessels, holding both the company's DOC certificate and SMC certificates for individual ships in effect by an International Safety Management Code (ISM), and had obtained ISO certificates on the quality management system. A month prior, an internal audit of the vessel was conducted by the company, and two identified non-conformities had been rectified. There was a record of supply of all items on time which were requested by the ship. A total of 22 crew members were on board and were made up of two different Nationalities. All crew members who had previously participated as rescue crews had relevant

certificates as qualified rescue crew. The captain had more than twenty years of experience and had been working for the company for 5 years as a captain and had been working aboard the vessel since the time of ship delivery. The chief officer had seven years of experience and had been onboard the ship for three months. The officer on duty boarded the ship the previous month, with three years of experience as a third officer. The watchkeeping crew had five years of experience as an AB. Crewman-A, the person in charge of rescue boat control, had seven years of sailing experience as a second officer and had also been on board since ship delivery. Crewman-B and –C had boarded three months ago with three years of experience as oilers. Crewman-D, who was responsible for winch control, boarded six months ago with a twenty-year career as a bosun. As a result of the health record review and interview of all crew members, there were currently no crew members taking medications or experiencing physical and/or mental ill-health. The muster list, including procedures for rescuing men overboard, was posted on the walls of the navigation bridge, each corridor and cafeteria, as well as on their respective duty pocketbook. The search and rescue procedure, including William's turn, is posted on the wall of the navigation bridge. The rescue boat davit posts how it works with an illustration. Instruction on the operation of the release hook within the rescue boat is posted with the illustration.

The ship's captain planned to conduct the man overboard drill with a No.1 enclosed type lifeboat, which is assigned as a rescue boat, at 3 p.m. Once the drill begins, the watchkeeping crew on duty identify a man overboard and throw a life penalty. The lifebuoy is a quick-release type, which automatically drops when the safety pin is removed, and the lever is pulled, and report to the officer on duty. The officer on duty reduces the vessel speed, marks the man's location falling into the sea, immediately reports to the captain and the relevant authorities, and uses a public address system and alarm to notify other crew members. As soon as the captain is aware of the situation, they move to the navigation bridge and manoeuvre the ship back to the location where the person was reported as overboard. The rest of the crew gather with personal equipment at the muster station on the boat deck. The chief officer instructs the wearing of personal equipment and performs the inspection. The chief officer then delivers the safety briefing to the crew and instructs them to prepare for the launch of the rescue boat. These tasks include disconnecting the charging cable socket from the boat, removing the securing wire and lashing stopper, and connecting a painter line to hook on the rescue boat FWD. Crewman-A, -B, and –C are designated as rescue crews and board the rescue boat, while the rest of the crew are responsible for helping to launch and recover a rescue boat by controlling a davit winch. The rescue boat is a davit launching type and is lowered by gravity when the winch brake is released. The winch breaker is operated using a lever after removing the safety pin. The boat is raised by pressing a button on the remote controller to operate the winch. Crewman-D is responsible for winch operation. The boat release hook system can be operated in both on-load and off-load conditions. A hydrostatic interlock device is installed to prevent crew members from falling out before reaching sea level. The hydrostatic interlock can be manually released if the water pressure is not working properly due to severe sea conditions even though the boat has reached water level. In this case, releasing is called on-load release, while scenarios where the boat is buoyant and released without force applied to the hook, are deemed an off-load release. The release hook is operated using a lever after removing the safety pin by Crewman-A. Then, the pull lever located on the rescue boat's FWD is pulled, and the painter line is removed before rescue operations begin in earnest. Once the rescue boat is completely removed from the primary vessel, the rescue boat is manoeuvred to begin rescue operations. At the end of the rescue activity, a suspension link from the davit fall is connected to the hook of the rescue boat for boat recovery. If the rescue boat is properly connected to the wire of the davit, the winch is activated to raise the boat to the stowed position. More specific tasks are described in the task analysis.

4.2. Task analysis

The hierarchical task analysis for the procedures of the rescue boat drill for a man overboard is shown in Fig. 4. The procedure consists of

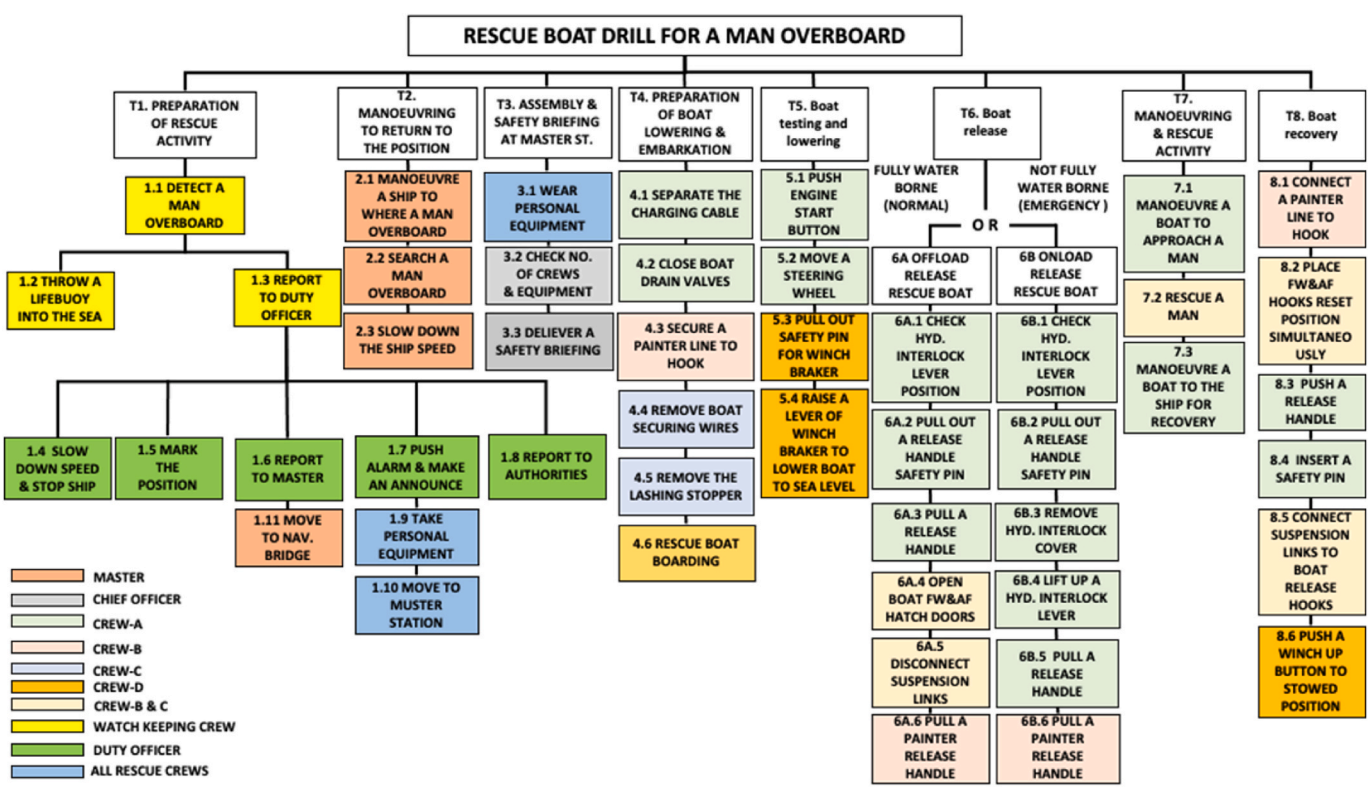

Fig. 4. Hierachical task analysis.

Table 3
Tabular task analysis (part of).

Task	Responsible person	Location	H-M interface	Equipment	Required additional manual
1.1 Detect a man overboard	Watchkeeping crew	Wing bridge	N/A	N/A	N/A
1.2 Throw a lifebuoy into the sea	Watchkeeping crew	Wing bridge	Safety pin, lever handle	Quick-release lifebuoy	Manual for quick release lifebuoy
1.3 Report to the duty officer	Watchkeeping crew	Nav. bridge	N/A	N/A	N/A
1.4 Slow down speed and stop the ship	Duty officer	Nav. bridge	Lever handle	Engine telegraph	N/A
1.5 Mark the position where a man overboard	Duty officer	Nav. bridge	Display screen	Chart or ECDIS	N/A
1.6 Report to Master	Duty officer	Nav. bridge	Telephone	Telephone	N/A
1.7 Push alarm and make an announcement	Duty officer	Nav. bridge	Push-button, Announce device	Alarm system, P.A. system	Manual for alarm system, Manual for PA system
1.8 Report to Authorities	Duty officer	Nav. bridge	VHF	VHF	Manual for radio equipment and contact details
1.9 Take personal equipment	All rescue crews	Cabin room	N/A	Personal equipment	N/A

eight main tasks, which are i) Preparation of rescue activity, ii) Ship manoeuvring, iii) Assembly and safety briefing, iv) Preparation of boat lowering and embarkation, v) Boat testing and lowering, vi) Boat release, vii) Manoeuvring and rescue activity, and viii) Boat recovery. Additional information obtained from the Tabular task analysis is described in Table 3.

4.3. Assessment of performance shaping factors

Five maritime experts are carefully selected for this assessment, where experts are asked to select a linguistic scale of PSFs for each task. Then, qualitative expert opinions are aggregated. The relative importance among experts is considered as a heterogeneous group, depending on their background. A relaxation factor (β) is assumed to be 0.4, and relative importance w_i among experts is determined as 0.20, 0.18, 0.21, 0.20 and 0.21 for five experts. As an example, specific opinion aggregation for Task 8.2 is illustrated in Table 4 and Table 5. Once experts' judgment and fuzzy opinion aggregation are completed, the aggregated rating is converted to the multiplier for each PSF for a human error quantification, following equations in Table 1.

4.4. Human error quantification

The human error probabilities for each sub-task of a man overboard procedure during the whole rescue boat drill is listed in Table 6. Human error probability is computed based on the SPAR-H human quantification technique, which is described in Equations (2), (3) and (12).

4.5. Human reliability assessment

Once human error probabilities for each sub-task are derived, the final step is to incorporate the human error probability of sub-task into Hierarchical task analysis in Fig. 4 to derive single failure probability for overall assessment by considering system description and dependency of sub-tasks based on the Rules in Table 2. However, as mentioned in section 3.5, to apply these Rules to each major task, the sub-tasks should be sequentially connected for evaluation of dependency, and the tasks should also be decomposed to the level to which the same system can be assumed. This section introduces the following techniques for converting from HTA to reliability block diagrams for each task with different characteristics. According to the HTA in Fig. 4, the entire procedure consists of eight major tasks, and each task consists of each sub-task. Task 1 consists of eleven sub-tasks that must be successfully performed to complete the preparation of rescue activities. However, sub-tasks for task 1 do not always occur sequentially, and some sub-tasks are linked to multiple sub-tasks. This makes it challenging to connect sub-tasks to either serial or parallel systems for task 1. For example, sub-task 1.1 in Fig. 4 requires actions in sub-task 1.2 and 1.3, and sub-task 1.3 initiates five actions in sub-task 1.4, 1.5, 1.6, 1.7, and 1.8, which may occur in any order. For human reliability modelling, task 1 is decomposed by six different groups of sub-tasks for missions, and each

Table 5
Opinion aggregation working condition of task 8.2

Degree of agreement(S)		The relative degree of agreement (RA)	
S12	0.86	RA(Ex1)	0.21
S23	0.86	RA(Ex2)	0.20
S13	1.00	RA(Ex3)	0.21
S14	0.71	RA(Ex4)	0.18
S24	0.86	RA(Ex5)	0.21
S34	0.71		
S15	1.00	Consensus degree coefficient (CC)	
S25	0.86	CC(Ex1)	0.20
S35	1.00	CC(Ex2)	0.19
S45	0.71	CC(Ex3)	0.21
The average degree of agreement (AA)		CC(EX4)	0.19
AA(Ex1)	0.89	CC(Ex5)	0.21
AA(Ex2)	0.86		
AA(Ex3)	0.89	Result of aggregation (Rag)	
AA(Ex4)	0.75		(2.44, 3.44, 4.44)
AA(Ex5)	0.89	Defuzzified rating	3.44

Table 4
Experts' evaluation for PSFs of task 8.2

PSFs	PSF1	PSF2	PSF3	PSF4	PSF5	PSF6	PSF7	PSF8	PSF9
Expert 1	N	N	N	N	N	LN	LN	VLN	N
Expert 2	VLN	N	N	VLN	N	VLN	LN	N	LP
Expert 3	N	N	N	VLN	N	LN	VLN	VLN	N
Expert 4	N	VLN	N	VLN	N	N	N	N	LP
Expert 5	N	VLN	N	VLN	N	LN	LN	N	N
Aggregated rating	2.18	2.39	2.00	2.81	2.00	3.44	3.42	2.40	1.61
Interpolated multiplier	1.18	1.39	1.00	1.81	1.00	3.32	3.27	1.40	0.81

*N is nominal, VLN is very low negative, LN is Low negative, LP is Low positive.

Table 6
Human error probability for each sub-task of a man overboard procedure.

Task	Diagnosis Error	Execution Error	Human Error
1. Preparation of rescue activity			
1.1 Detect a man overboard	2.04E-02	2.04E-03	2.24E-02
1.2 Throw a lifebuoy into the sea	5.72E-03	5.72E-04	6.29E-03
1.3 Report to the duty officer	3.60E-03	3.60E-04	3.96E-03
1.4 Slow down speed and stop the ship	3.30E-03	3.30E-04	3.63E-03
1.5 Mark the position where a man overboard	3.05E-03	3.05E-04	3.36E-03
1.6 Report to Master	2.23E-03	2.23E-04	2.45E-03
1.7 Push alarm and make an announcement	2.03E-03	2.03E-04	2.23E-03
1.8 Report to Authorities	9.81E-03	9.81E-04	1.08E-02
1.9 Take personal equipment	2.91E-03	2.91E-04	3.20E-03
1.10 Move to muster station	3.65E-03	3.65E-04	4.01E-03
1.11 Move to Navigation bridge	3.65E-03	3.65E-04	4.01E-03
2. Ship manoeuvring			
2.1 Manoeuvring a ship to where a man overboard	1.59E-02	1.59E-03	1.75E-02
2.2 Search a man overboard	2.08E-02	2.08E-03	2.28E-02
2.3 Slow down the ship speed	1.58E-03	1.58E-04	1.74E-03
3. Assembly and safety briefing			
3.1 Wear personal equipment	3.33E-03	3.33E-04	3.66E-03
3.2 Check number of crews and their equipment	2.84E-03	2.84E-04	3.12E-03
3.3 Deliver a safety briefing	9.52E-03	9.52E-04	1.05E-02
4. Preparation of boat lowering and embarkation			
4.1 Separate the charging cable	6.47E-03	6.47E-04	7.11E-03
4.2 Close boat drain valves	1.12E-02	1.12E-03	1.23E-02
4.3 Secure a painter line to the rescue boat painter hook	1.70E-02	1.70E-03	1.86E-02
4.4 Remove boat securing wires	1.33E-02	1.33E-03	1.47E-02
4.5 Remove the lashing stopper on boat davit	1.33E-02	1.33E-03	1.46E-02
4.6 Rescue boat boarding	8.24E-03	8.24E-04	9.06E-03
5. Boat testing and lowering			
5.1 Push the engine start button	1.04E-02	1.04E-03	1.15E-02
5.2 Move the steering wheel	8.93E-03	8.93E-04	9.82E-03
5.3 Pull out the safety pin for the winch brake	5.45E-03	5.45E-04	6.00E-03
5.4 Raise the lever of the winch brake to lower the boat to the sea level	1.30E-02	1.30E-03	1.43E-02
6. Boat release			
6A Off-load release rescue boat			
6A.1Check hydrostatic interlock lever position	1.56E+01	1.56E-01	1.56E-02
6A.2 pull out the release handle safety pin	1.04E+01	1.04E-01	1.04E-02
6A.3 Pull the release hand	1.04E+01	1.04E-01	1.04E-02
6A.4 Open the boat F & A hatch doors	5.90E+00	5.90E-02	5.90E-03
6A.5 Disconnect suspension links from hook	2.19E+01	2.19E-01	2.19E-02
6A.6 Pull the painter release handle	1.57E+01	1.57E-01	1.57E-02
6B On- load release rescue boat			
6B.1Check hydrostatic interlock lever position	1.56E+01	1.56E-01	1.56E-02
6B.2 Pull out the release handle safety pin	1.04E+01	1.04E-01	1.04E-02
6B.3 Remove the hydrostatic interlock cover	1.29E+01	1.29E-01	1.29E-02
6B.4 Lift up the hydrostatic interlock lever	1.03E+01	1.03E-01	1.03E-02
6B.5 Pull the release handle	1.04E+01	1.04E-01	1.04E-02
6B.6 Pull the painter release handle	1.57E+01	1.57E-01	1.57E-02
7. Manoeuvring and rescue activity			
7.1 Manoeuvring the rescue boat to approach a man	2.49E-01	2.49E-02	2.68E-01
7.2 Rescue a man (Pull up a man to a boat)	2.65E-01	2.65E-02	2.85E-01
7.3 Manoeuvring the rescue boat to the ship for recovery	2.09E-01	2.09E-02	2.26E-01
8. Boat recovery			
8.1 Connect a painter line to the painter hook	1.67E-01	1.67E-02	1.81E-01
	3.65E-01	3.65E-02	3.89E-01
8.2 Place FWD & AFT hooks reset position simultaneously			
8.3 Push the release handle	8.98E-02	8.98E-03	9.80E-02
8.4 Insert the release handle safety pin	9.02E-02	9.02E-03	9.84E-02
8.5 Connect suspension links to boat release hooks	2.10E-01	2.10E-02	2.26E-01
8.6 Push the winch up button to stowed position	7.03E-03	7.03E-04	7.73E-03

group is assessed by their system configuration and dependency of sub-tasks, as shown in Fig. 5. Specifically, the reliability of group R1-1 is assigned as the maximum value of two sub-tasks because sub-tasks 1.1 and 1.2 are high dependency in the serial system. That is, if one of the two sub-works fails, the mission for group R1-1 fails, and the success or failure of preceding sub-task 1 affects the conditional probability of sub-task 1.2. Similarly, missions from group R1-2 to R1-6 are configured, as shown in Fig. 6. Once reliabilities for all mission from group R1-1 to R1-6 are derived, total reliability for task 1 is assigned as multiplication of each group's reliabilities and the value is 8.83E-01 because if any of six missions fails the task 1 will fail (series system) and six missions have low dependency. For task 2, reliability for task 2 is assigned as minimum reliability since sub-tasks from 2.1 to 2.3 are in the series system, and three subtasks are high dependency. For task 3, sub-task 3.2 'Check number of crews and their equipment is redundancy for the sub-task3.1 'Wear personal equipment relation. It means any of the two sub-tasks succeed, task 3 will succeed. Therefore, group reliability for sub-task 3.1 & 3.2 is calculated by equation (16) because they are in the parallel system with low dependency. The total reliability for task 3 is assigned as multiplication of reliabilities for a group of sub-tasks 3.1 & 3.2 and sub-task 3.3 since they are in a series system with low dependency and the value is 9.89E-01 as shown in Fig. 7. In task 4, the sub-task from 4.1 to 4.6 is configured in a series system with no dependence. This means that task 4 succeeds only when all sub-works are successfully completed, but the success or failure of each sub-task does not affect the success or failure of other sub-tasks. For example, sub-task 4.1 'Separate the charging cable 'and sub-task 4.2 'Close boat drain vales 'should be completed successfully before rescue boat launching, but each sub-tasks 4.1 and 4.2 does not affect each other's result. The reliability of task 4 is the product of the reliabilities of individual sub-tasks. For task 5, four sub-tasks are classified as two different series systems with high dependency. Then two groups are combined in a series system with no dependency, as shown in Fig. 7. The value of the reliability of task 5 is 9.74E-01. For tasks 6A, 6B, 7 are all series systems with high dependency. Therefore, the minimum reliability of their sub-tasks is assigned as each task's reliability. The values are 9.78E-01, 9.84E-01 and 7.15E-01, respectively. Similarly, the reliability of task 8 is assigned as 6.016E-01 in Fig. 7. Finally, the total reliability of the rescue drill scenario can be derived from the reliability block diagram in Fig. 8 by computing the reliabilities of each task. In order to determine the final failure probability value for man overboard drill, eight main tasks should succeed individually. These tasks should be conducted sequentially and highly dependent. Therefore, the minimum reliability of eight tasks is assigned as overall reliability for the whole procedure, and the value is 6.06E-01.

4.6. Findings and discussion

The proposed approach presents individual human error probabilities obtained by a proposed method based on a particular maritime scenario: a rescue boat drill for a man overboard procedure. The overall high human error has been demonstrated during the manoeuvring and rescue activity of task 7. This seems to have adversely affected the human performance as the narrow space in the enclosed boat restricted

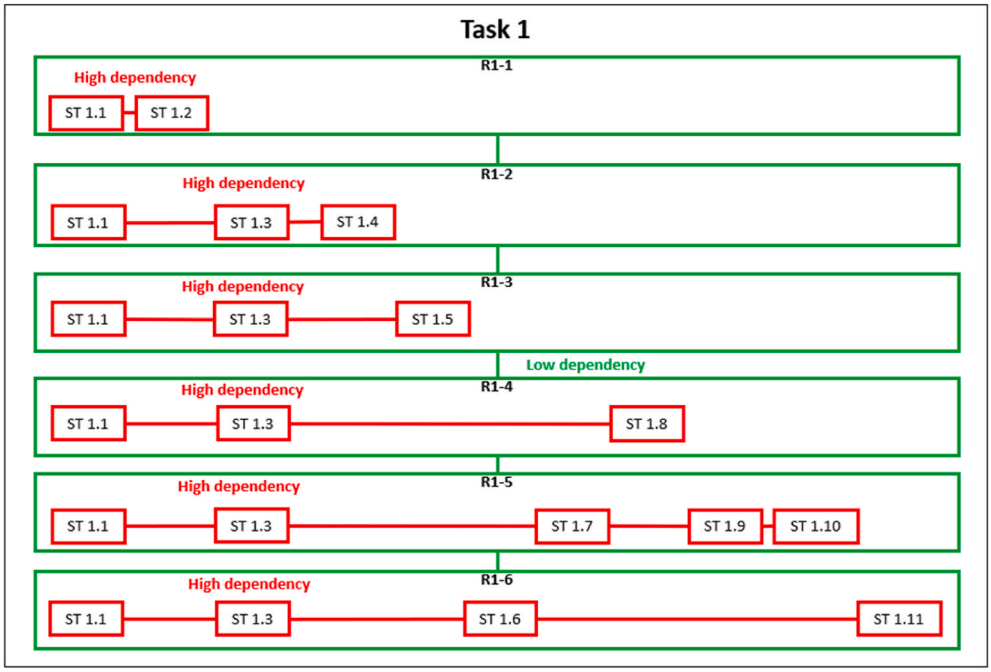

Fig. 5. Sub-task level reliability block diagram for task 1.

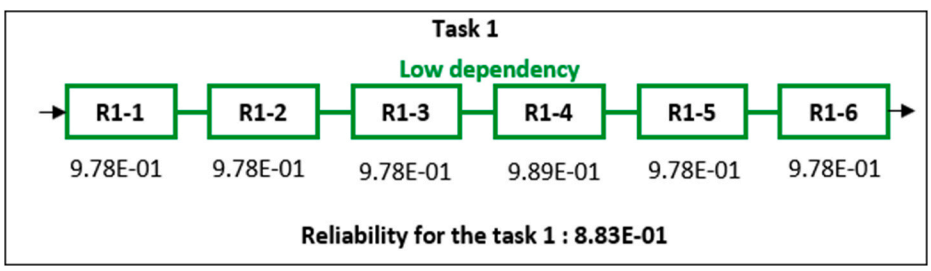

Fig. 6. Reliability block diagram for task 1.

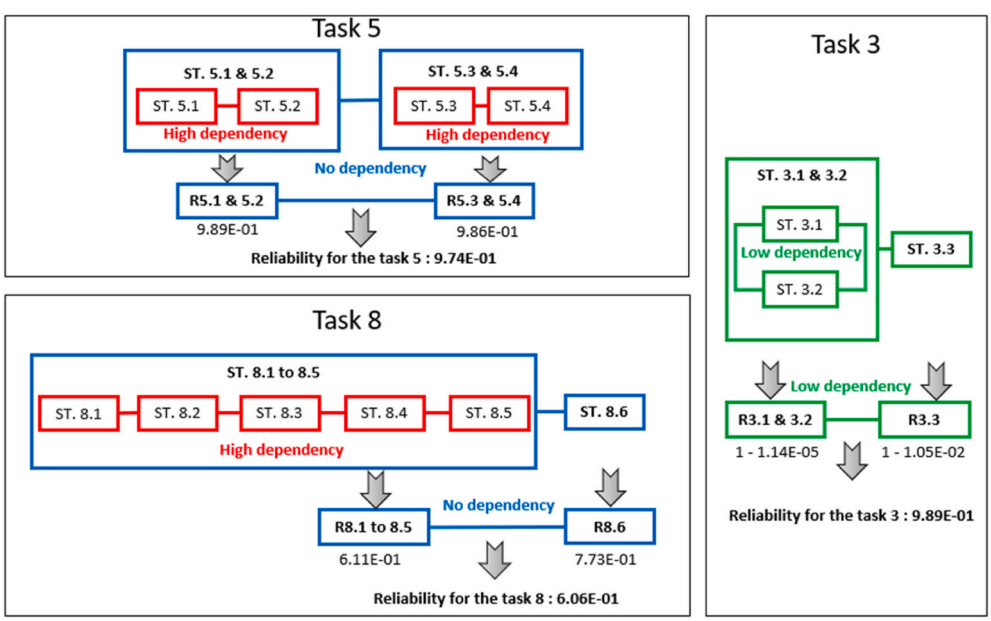

Fig. 7. Reliability block diagram for a different types of tasks.

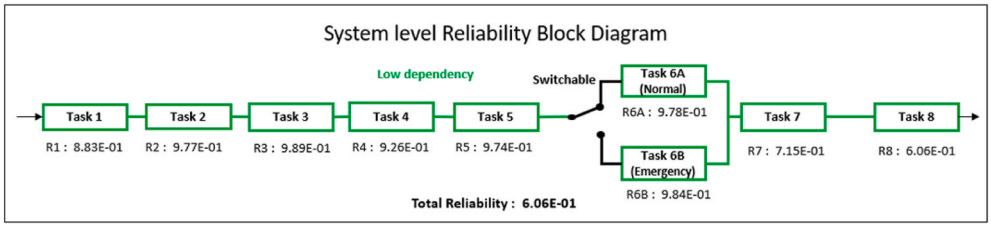

Fig. 8. System-level reliability block diagram.

the crew's free movement, affected the crew's vision, and even the high level of noise interrupted communication. Additionally, waves and currents at sea affected the main vessel, making it difficult to control the ship. Therefore, these ambient factors increase stress exerted on the crew and pressure on time for direct rescue activities. Sub-task 8.2 (Place FWD & AFT hooks reset position simultaneously) in the boat recovery (task 8) has the highest HEP value as of 3.89E-01 during the whole procedure, due to the unfavourable circumstances (such as the manoeuvring and rescue activity phase) and each crew member should perform their task simultaneously, even if it is a simple task. The second-highest human error probability is that of sub-task 7.2 (Rescue a man) at 2.85E-01. Sub-task 7.1 (Manoeuvring the rescue boat to approach a man) at 2.68E-01; sub-task 7.3(Manoeuvring the rescue boat to the ship for recovery) & sub-task 8.5 (Connect suspension links to boat release hooks) are 2.26E-01 in order, are also at notably high probabilities of failure. Conversely, sub-tasks 2.3 (Slow down the ship speed), 1.6 (Report to Master) and 1.7 (Push alarm and make an announcement) show the lowest HEP with a range from 1.74E-03 to 2.45E-03. In human reliability assessment for each task, tasks 1 through 4 are to prepare for rescue operations where human reliabilities intervals range from 8.83E-01 to 9.89E-01. The operations in the rescue boat are divisible: task 5 for embarkation at the boat stowed position; task 6 for boat release where the boat is hung on the wire fall; task 7 for rescue activities; and task 8 for boat recovery at sea condition. The human reliability ranges for tasks 5 through 8 in a rescue boat operation are in 6.06E-01 to 9.84E-01. These different human reliability intervals, occurring in the same rescue boat and while performing similar tasks, demonstrate that even if the task takes place in the same workspace with the same crew, human performance changes depending on the characteristics of the circumstance. The task with the lowest reliability is the process of boat recovery of task 8. Compared to task 6 (Release rescue boat) 's reliability interval of 9.78E-01 to 9.8.4E-01 depending on the type of boat release, the boat recovery process causes more human errors than release.

Interestingly, whether the rescue boat is released on-load or off-load condition, it does not change the reliability of the entire process. This means tasks that take place on sea conditions after rescue boat released from wire fall significantly affect whole reliability. Finally, the PSFs contributing to human error in rescue scenarios for each task phase are illustrated in Fig. 9. However, the extent to which PSF affects human performance may vary from sub-task to sub-task, even within the same task. The most significant contributing factors are the ambient conditions for the workplace and environmental conditions in tasks 6 to 8 related to the rescue boat operation at sea, which is at the core of the procedures.

Unlike the use of monolingual variables in conventional SPAR-H, fuzzy sets were adopted to explain better the effects of PSF, which provide a multiplier corresponding to a continuous number obtained using fuzzy and their interpolated values rather than integers, thereby increasing. Furthermore, for integrating this dependency information, the proposed RBD provide a relatively simple process to address dependency than the THERP method, which require a large amount of effort to produce reliable HEP. Furthermore, the proposed method also provides a unique system structure to calculate the effect of recovery action.

5. Conclusion

Human error is one of the main contributing factors in failure during emergency preparedness in maritime transportation. Although human error assessment is a critical issue in maritime safety, the quantification process is difficult due to the limited human error data. Therefore, many parts of the quantisation process rely on the qualitative judgment of experts. It is important to form consensus properly and to effectively quantify the collected diverse opinions. In this context, this paper

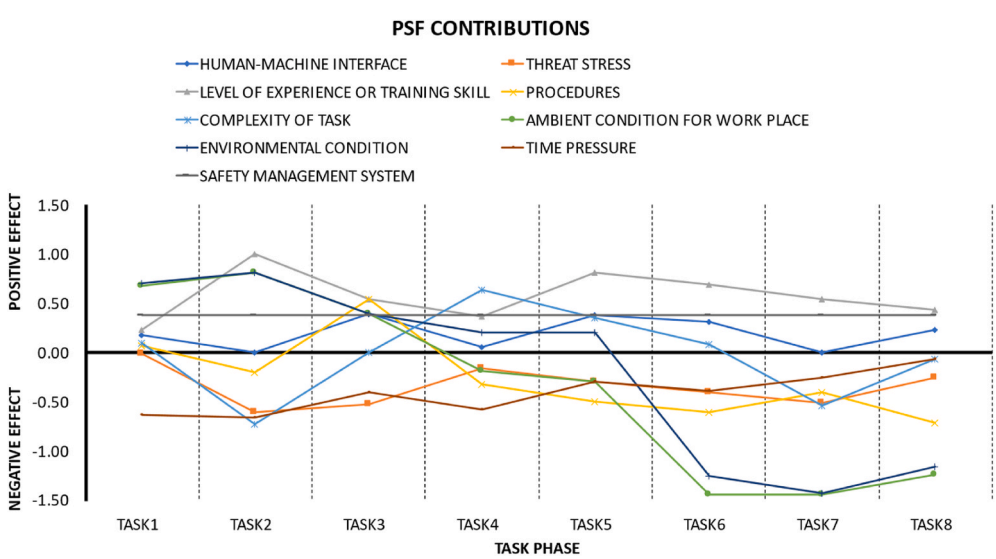

Fig. 9. PSFs contribution for task phases.

introduces a new framework-based SPAR-H approach applicable to maritime emergency drill scenarios and illustrates practical rescue boat drill procedures for a man overboard. There are various characteristics and expected advantages of the proposed method. Firstly, the proposed method provides refined PSFs with customised guidance to reflect on-board rescue drills' characteristics for expert opinion collection. The selected PSFs have a language variable representing the level of PSF that deals negatively or positively with the expected impact on performance reliability and developed a fuzzy set to describe the impact of PSF better. Secondly, the fuzzy opinion aggregation method converts experts' multiple qualitative assessments of PSF ratings into one integrated opinion with fuzzy opinions and converts them into crisp values through defuzzification. SPAR-H then calculates human error probabilities based on PSF ratings from an expert's opinion aggregation method for tasks during rescue boat drills. This hybrid approach may enhance the reliability and consistency of the outcomes. Notably, the novel approach to model a human reliability assessment from individual human error probabilities based on a reliability block diagram is applicable to various types of systems. In addition, through this approach, the relationship between complex tasks is effectively displayed in a simplified way which cannot be achieved solely by applying hierarchical task analysis.

One of the limitations of this method, as already mentioned, is the fact that the technique is developed for the nuclear industry. Therefore, it may not be fully applicable to all situations of maritime operation. Therefore, more research is needed to develop and verify customised PSFs for more diverse working conditions in the maritime industry. In addition, the presented framework does not fully consider complex interactions between humans, machines and software because it is more difficult to identify human errors in complex systems. As a result of adoption of new technologies, it will be more difficult to identify human functions in a complex system with traditional approaches. Therefore, more powerful approaches to identify human error is needed.

In conclusion, maintaining emergency preparedness is undeniably essential in ship operation. However, it is also important to evaluate whether training for emergency preparedness is in a suitable state to be implemented. This study provides a framework for conducting a human reliability assessment of emergency training, which can help ship operators in the decision-making process and have a positive impact on the safety of ship operations and maritime safety.

CRediT authorship contribution statement

Sung Il Ahn: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization, Project administration, Data curation. **Rafet Emek Kurt:** Conceptualization, Validation, Resources, Writing – review & editing, Supervision, Project administration. **Emre Akyuz:** Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abbassi, R., Khan, F., Garaniya, V., Chai, S., Chin, C., Hossain, K.A., 2015. An integrated method for human error probability assessment during the maintenance of offshore facilities. Process Saf. Environ. Protect. 94, 172–179.

Ahn, S.I., Kurt, R.E., 2020. Application of a CREAM based framework to assess human reliability in emergency response to engine room fires on ships. Ocean Eng. 216, 108078.

Akyuz, E., 2015. Quantification of human error probability towards the gas inerting process on-board crude oil tankers. Saf. Sci. 80, 77–86.

Akyuz, E., 2016. Quantitative human error assessment during abandon ship procedures in maritime transportation. Ocean Eng. 120, 21–29.

Akyuz, E., Celik, E., Celik, M., 2018. A practical application of human reliability assessment for operating procedures of the emergency fire pump at ship. Ships Offshore Struct. 13, 208–216.

Akyuz, E., Celik, M., 2014. Utilisation of cognitive map in modelling human error in marine accident analysis and prevention. Saf. Sci. 70, 19–28.

Akyuz, E., Celik, M., 2015. Application of CREAM human reliability model to cargo loading process of LPG tankers. J. Loss Prev. Process. Ind. 34, 39–48.

Akyuz, E., Celik, M., 2016. A hybrid human error probability determination approach: the case of cargo loading operation in oil/chemical tanker ship. J. Loss Prev. Process. Ind. 43, 424–431.

Antåo, P., Soares, C.G., 2019. Analysis of the influence of human errors on the occurrence of coastal ship accidents in different wave conditions using Bayesian Belief Networks. Accid. Anal. Prev. 133, 105262.

Bell, J., Williams, J., 2016. Consolidation of the HEART Human Reliability Assessment Principles. Symposium Series No.

Blackman, H.S., Gertman, D.I., Boring, R.L., 2008. Human error quantification using performance shaping factors in the SPAR-H method. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications Sage CA, Los Angeles, CA, pp. 1733–1737.

Bye, A., Laumann, K., Taylor, C., Rasmussen, M., Øie, S., Van De Merwe, K., Øien, K., Boring, R., Paltrinieri, N., WèRú, I., 2017. The Petro-HRA Guideline.

Erdem, P., Akyuz, E., 2021. An interval type-2 fuzzy SLIM approach to predict human error in maritime transportation. Ocean Eng. 232, 109161.

Fujita, Y., Hollnagel, E., 2004. Failures without errors: quantification of context in HRA. Reliab. Eng. Syst. Saf. 83, 145–151.

Groth, K.M., Smith, C.L., Swiler, L.P., 2014. A Bayesian method for using simulator data to enhance human error probabilities assigned by existing HRA methods. Reliab. Eng. Syst. Saf. 128, 32–40.

Groth, K.M., Swiler, L.P., 2013. Bridging the gap between HRA research and HRA practice: a Bayesian network version of SPAR-H. Reliab. Eng. Syst. Saf. 115, 33–42.

He, X., Wang, Y., Shen, Z., Huang, X., 2008. A simplified CREAM prospective quantification process and its application. Reliab. Eng. Syst. Saf. 93, 298–306.

Hollnagel, E., 1998. Cognitive Reliability and Error Analysis Method (CREAM). Elsevier.

Imo, S.C.E., 2018. IMO. London.

Islam, R., Abbassi, R., Garaniya, V., Khan, F., 2017a. Development of a human reliability assessment technique for the maintenance procedures of marine and offshore operations. J. Loss Prev. Process. Ind. 50, 416–428.

Islam, R., Yu, H., Abbassi, R., Garaniya, V., Khan, F., 2017b. Development of a monograph for human error likelihood assessment in marine operations. Saf. Sci. 91, 33–39.

Jahangiri, M., Hoboubi, N., Rostamabadi, A., Keshavarzi, S., Hosseini, A.A., 2016. Human error analysis in a permit to work system: a case study in a chemical plant. Safety and health at work 7, 6–11.

Khan, F., 2008. Human factors special issue. J. Loss Prev. Process. Ind. 21, 225–226.

Kirwan, B., 1994. A Guide to Practical Human Reliability Assessment. CRC press.

Kirwan, B., Ainsworth, L.K., 1992. A Guide to Task Analysis: the Task Analysis Working Group. CRC press.

Konstandinidou, M., Nivolianitou, Z., Kiranoudis, C., Markatos, N., 2006. A fuzzy modeling application of CREAM methodology for human reliability analysis. Reliab. Eng. Syst. Saf. 91, 706–716.

Kristiansen, S., 2013. Maritime Transportation: Safety Management and Risk Analysis. Routledge.

Kurt, R.E., Khalid, H., Turan, O., Houben, M., Bos, J., Helvacioglu, I.H., 2016. Towards human-oriented norms: considering the effects of noise exposure on board ships. Ocean Eng. 120, 101–107.

Lee, S.M., Ha, J.S., Seong, P.H., 2011. CREAM-based communication error analysis method (CEAM) for nuclear power plant operators' communication. J. Loss Prev. Process. Ind. 24, 90–97.

Luo, M., Shin, S.-H., 2019. Half-century research developments in maritime accidents: future directions. Accid. Anal. Prev. 123, 448–460.

Mirzaei Aliabadi, M., Esmaeili, R., Mohammadfam, I., Ashrafi, M., 2019. Human reliability analysis (HRA) using standardized plant analysis risk-human (SPAR-H) and bayesian network (BN) for pipeline inspection gauges (PIG) operation: a case study in a gas transmission plant. Health Scope 8.

Navas De Maya, B., Ahn, S., Kurt, R., 2019. Statistical Analysis of MAIB Database for the Period 1990-2016.

Noroozi, A., Khan, F., Mackinnon, S., Amyotte, P., Deacon, T., 2014. Determination of human error probabilities in maintenance procedures of a pump. Process Saf. Environ. Protect. 92, 131–141.

Olcer, A., Odabasi, A., 2005. A new fuzzy multiple attributive group decision making methodology and its application to propulsion/manoeuvring system selection problem. Eur. J. Oper. Res. 166, 93–114.

Parhizkar, T., Utne, I.B., Vinnem, J.E., Mosleh, A., 2021. Dynamic probabilistic risk assessment of decision-making in emergencies for complex systems, case study: dynamic positioning drilling unit. Ocean Eng. 237, 109653.

Reason, J., 1990. Human Error. Cambridge university press.

Ross, T.W., 2006. Ship's Lifeboats: Analysis of Accident Cause and Effect and its Relationship to Seafarers' Hazard Perception. Master's, Dalhousie University. Retrieved from. http://www.msc1206.com.

Salas, E., Driskell, J.E., Hughes, S., 1996. The Study of Stress and Human Performance. *Stress and Human Performance(A 97-27090 06-53)*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 1–45. Publishers, 1996.

Schrøder-Hinrichs, J.U., Baldauf, M., Ghirxi, K.T., 2011. Accident investigation reporting deficiencies related to organizational factors in machinery space fires and explosions. Accid. Anal. Prev. 43, 1187–1196.

Sezer, S.I., Akyuz, E., Arslan, O., 2022. An extended HEART Dempster–Shafer evidence theory approach to assess human reliability for the gas freeing process on chemical tankers. R*eliab. Eng. Syst. Saf.* 220, 108275.

Shappell, S.A., Wiegmann, D.A., 2000. The Human Factors Analysis and Classification System–HFACS.
Uflaz, E., Celik, E., Aydin, M., Erdem, P., Akyuz, E., Arslan, O., Kurt, R.E., Turan, O., 2022. An extended human reliability analysing under fuzzy logic environment for ship navigation. Aust. J. Maritime Ocean Affairs 1–21.
Ung, S.-T., 2015. A weighted CREAM model for maritime human reliability analysis. Saf. Sci. 72, 144–152.
Ung, S.-T., 2019. Evaluation of human error contribution to oil tanker collision using fault tree analysis and modified fuzzy Bayesian Network based CREAM. Ocean Eng. 179, 159–172.
Ung, S.T., Shen, W.M., 2011. A novel human error probability assessment using fuzzy modeling. Risk Anal.: Int. J. 31, 745–757.
Whaley, A.M., Kelly, D.L., Boring, R.L., Galyean, W.J., 2011. SPAR-H step-by-step guidance. In: Idaho Falls (ID): Risk, Reliability, and NRC Programs Department. Idaho National Laboratory.
Xi, Y., Yang, Z., Fang, Q., Chen, W., Wang, J., 2017. A new hybrid approach to human error probability quantification–applications in maritime operations. Ocean Eng. 138, 45–54.
Yang, Z., Bonsall, S., Wall, A., Wang, J., Usman, M., 2013. A modified CREAM to human reliability quantification in marine engineering. Ocean Eng. 58, 293–303.
Zhang, M., Zhang, D., Goerlandt, F., Yan, X., Kujala, P., 2019. Use of HFACS and fault tree model for collision risk factors analysis of icebreaker assistance in ice-covered waters. Saf. Sci. 111, 128–143.
Zhou, Q., Wong, Y.D., Xu, H., Van Thai, V., Loh, H.S., Yuen, K.F., 2017a. An enhanced CREAM with stakeholder-graded protocols for tanker shipping safety application. Saf. Sci. 95, 140–147.
Zhou, T., Wu, C., Zhang, J., Zhang, D., 2017b. Incorporating CREAM and MCS into fault tree analysis of LNG carrier spill accidents. Saf. Sci. 96, 183–191.