

Alarm Management at Sea

- An inquiry into time-critical actions on ships

We establish what it means for an alarm to be 'actionable', 'timely', and 'reasonable' in a way that enables objective evaluation as to whether a ship's general layout and system design allow adequate time for its engineering crew to carry out time-critical actions that fall within the range of empirically validated response times. Using the methodology outlined in ANSI/ANS 58.8-1994 Appendix 1.6, we establish time intervals for timely operator responses at high confidence levels. We then evaluate the applicability of ANSI/ANS 58.8-2019(RE2024) for verification of adequate time margins during preoperational and periodic validations on ships. Our contribution provides a basis for determining:

- (1) whether manual actions in response to a specific alarm can be credited within a ship's certification basis, or must be accomplished by an automatic action, or
- (2) if design modifications can obviate the need for such actions altogether, as well as
- (3) determine when credit taken for specific manual actions is reasonable.

Authors: Asger Schliemann-Haug (LR) (corresponding author), Peter Brown (LR), Flemming H. Pedersen (AAMS), Lars Anders Thomsen (AAMS), Stig Eriksen (AAMS), Niels Tingleff Haastrup (FMS), Casper Raahede(AAMS), Trine Heinemann (MARNAV), Martin Reimert Bach Pedersen (TEKNIKA), Jan Runge (TEKNIKA), Amelia Hipwell (MDH), Volkan Arslan (LR¹), Christopher Stein (RCG), Carlos De Sedas (RCG) Nanna Thit Hemmingsen (DFDS), Kristoffer Kloch (DFDS²).

(AAMS) - AARHUS SCHOOL OF MARINE AND TECHNICAL ENGINEERING;
(DFDS) - DFDS SEAWAYS;
(FMS) - FREDERICIA SCHOOL OF MARINE AND TECHNICAL ENGINEERING;
(LR) - LLOYDS REGISTER;
(MARNAV) – MARSTAL NAVIGATIONAL COLLEGE;
(MDH) – THE MARITIME DECARB HUB;
(RCG) - ROYAL CARRIBEAN GROUP;
(TEKNIKA) - COPENHAGEN COLLEGE OF MARINE ENGINEERING AND TECHNOLOGY MANAGEMENT;

Authors Contributions (CRediT):

Conceptualization: Asger Schliemann-Haug, Kristoffer Kloch, Nanna Thit Hemmingsen.

Data curation: Asger Schliemann-Haug, Trine Heinemann.

Formal analysis: Asger Schliemann-Haug, Flemming H. Pedersen, Lars Anders Thomsen, Stig Eriksen, Trine Heinemann.

Funding acquisition: Asger Schliemann-Haug, Kristoffer Kloch, Nanna Thit Hemmingsen, Casper Raahede, Niels Tingleff Haastrup.

Investigation: Asger Schliemann-Haug, Flemming H. Pedersen, Lars Anders Thomsen, Stig Eriksen, Niels Tingleff Haastrup, Martin Reimert Bach Pedersen, Jan Runge, Trine Heinemann, Christopher Stein, Carlos De Sedas, Amelia Hipwell, Volkan Arslan, Peter Brown.

Project administration: Asger Schliemann-Haug, Casper Raahede, Nanna Thit Hemmingsen.

Writing – original draft: Asger Schliemann-Haug.

Writing – review & editing: Niels Tingleff Haastrup, Kristoffer Kloch, Casper Raahede, Amelia Hipwell, Christopher Stein, Peter Brown.

Disclaimer: The contents of this document, including any opinions and conclusions expressed, is that of the authors alone and do not necessarily reflect the current position or policy of any of the authors associated organisations nor the opinion of any mentioned parties.

Intellectual property: This document includes descriptions of methods and designs for handling the management of alarms, which have been supplied by third parties. Readers should be aware that the intellectual property rights of these methods and designs may be held by those third parties or by others. Written permission was obtained before inclusion of such content in this document.

The authors thank the International Electrotechnical Commission, the American Nuclear Society, and the Engineering Equipment and Materials Users Association for permission to reproduce information from their standards and guidance. Neither organisation has any responsibility for the placement and context in which extracts and content are reproduced by the authors of this document, nor are these organisations in any way responsible for other content or accuracy therein.

¹ Now with CSMART - CENTER FOR MARITIME SIMULATION AND TRAINING

² Now with AAU - AALBORG UNIVERSITY

1	INTRODUCTION.....	4
2	Preliminaries.....	4
3	Industrial Boundaries.....	4
4	Motivation and Background.....	5
4.1	The context of engine room alarms.....	5
4.2	Losing control of alarms	5
5	Approach	6
5.1	Incident data	6
5.2	Naturalistic observations (field studies)...	6
5.3	Simulator studies	7
5.4	Statistical reduction	8
6	Relevant Good Practice	8
6.1	IEC 61511 - Functional safety for the processing industries	8
6.2	EEMUA 191 – Alarm systems, a guide to design, management and procurement.....	9
6.3	IEC 62682 – The management of alarm systems for the processing industries.....	9
6.4	NUREG 1278.....	10
6.5	ANSI/ANS 58.8.....	10
7	The method (Detailed).....	13
7.1	Discrete time points	13
7.2	Limiting design requirements	13
7.3	Discrete time intervals	14
7.4	Applications for multiple actions	14
7.5	Examples of applied methodology	15
8	Results and discussion.....	18
8.1	Acknowledgement time ($TI_{acknowledgement}$)	18
8.2	Diagnosis time ($TI_{diagnosis}$)	20
8.3	Operator errors and omissions.....	21
8.4	Operator time intervals ($TI_{dead} + TI_{operator}$)	21
8.5	Uncertainty and limitations of use	23
9	Conclusion and recommendations.....	24
9.1	Going forward	24
	Industry recommendations.....	24
10	Acknowledgments	25
11	Data availability statement	25
12	Works Cited	26

1 INTRODUCTION

Despite extensive investment in alarms during the construction of modern ships, little evidence exists as to whether the engineering crew can actually respond to them in time [1].

So do seafaring engineers have enough time to respond to alarms? and if not, what can we do to make sure they have going forward?

This work examines the time it takes for seafaring engineers to respond to abnormal situations (those that raise alarms). These response times include not only the crew's physical execution of corrective actions, but also the cognitive time taken for the *recognition, discrimination and evaluation* of the alarmed condition and *determination* of what action to take (diagnosis).

We argue that the likelihood of diagnosing an abnormal situation can be estimated on the basis of the available time to do so. Using observable actions alone, we present a nominal design model that accounts for the cognitive time delays in human alarm response. Moreover, we demonstrate that the model's objective estimates resonate closely with experts' subjective judgments.

Because an alarm is only effective if there is adequate time to act on it, the nominal model, based on the data gathered and presented in this work, reveals the time required for the crew to respond to abnormal situations at the 92.5% confidence level. This enables objective decisions on whether manual actions can be credited in response to alarms alerting the crew to *initiating events* that are foreseeable within a ship's design basis (design basis events: DBEs)³. Namely:

- 1 *whether manual actions can be relied upon to initiate or control a specific safety-related function or must be accomplished by automatic action, or*
- 2 *determine when design modifications can obviate the need for manual or automatic actions, and*
- 3 *if credit taken for manual actions is reasonable in cases where a single omission or error results in a consequence of concern.*

The purpose of this is to ensure that a ship's general layout and system design allow sufficient time for the crew to perform *time-critical actions* (TCAs) that fall within the range of empirically valid response times.

2 PRELIMINARIES

Alarm management is an established practice in industries adjacent to the maritime sector. Widely recognised standards and guidance on the subject, such as IEC 62682/ISA 18.2 and EEMUA 191, are based on the philosophy that all alarms should be justified (*rationalised*) to be worthy of the operators' time and effort for all system states and operational conditions in which they are presented [2] [3].

The IEC 62682/ISA 18.2 and EEMUA 191 all share a common definition of an alarm as "*an audible and/or visible means of indicating to the operator an equipment malfunction, process deviation, or abnormal condition requiring a timely response*"⁴.

In the maritime industry, the IMO Code on Alerts and Indicators define an alarm as a high-priority alert indicating a "*Condition requiring immediate attention and action to maintain the safe navigation and operation of the ship.*" - [4]

These definitions share a focus on the words *immediate* and *timely*, as well as *action* and *response*. The core intent is that if alerts are to be of any real use, they must at least be:

- 1 *Actionable. The alert indicates a condition that a human can do something about;*
- 2 *Timely. There is sufficient time to do what is necessary;*

At the same time, either definition implies *uniqueness*, as both the abnormal condition and the alarm are singular (there should not be several alarms for one single condition or vice versa).

The IMO is setting similar criteria for no less than seven distinct alert qualities in its forthcoming *Engine Control Room Alert Management* (ECRAM) performance standard [5, p. 6 cl. 29].

Readers of this document will have a better idea of how to determine whether a proposed alarm is *Actionable* and *Timely*. And do so preferably at a time when design changes can still be made at reasonable cost.

3 INDUSTRIAL BOUNDARIES

We present human performance data on responses to abnormal situations, both simulated and real. It is specific to engineering personnel on ships. Use outside this context constitutes extrapolation and should be treated with caution.

³ Initiating event: an event that is a condition of a device failure, system failure, external event, or wrong action (or inaction) that begins a sequence of events leading to a consequence of concern. Initiating events associated with one or more provisioned alarm points must be seen as anticipated and therefore deemed foreseeable in the ship design basis, i.e., a Design Basis Event (DBE), namely, an event for which the ship must be designed.

⁴ IEC 62682 ed.2.0 "Copyright © 2022 IEC Geneva, Switzerland. www.iec.ch

4 MOTIVATION AND BACKGROUND

The maritime industry relies on auditory and visual alerts to cue crew members to abnormal conditions that require their timely intervention. For this reason, the provision of brightly coloured and noisy alerts in the form of alarms constitutes a key part of the maritime safety regime.

4.1 The context of engine room alarms

A ship's machinery system is usually required to be designed as 'fail-safe' (usually means to switch off itself). In many aspects, this makes perfect sense, because it allows the crew to address the issue and resume operations. It would be undesirable for machinery to continue operating beyond the point of irreversible damage or degradation (we certainly don't desire to see ships stranded in the middle of the Atlantic because of a missed alarm). Therefore, many alarms are initially triggered when machinery reaches the safe state (tripped). This could be called a 'trip first, ask later' philosophy⁵.

At the same time, a ship is a (large) moving object. Here safety can sometimes mean keeping systems "on", rather than "off". Several high-profile incidents, such as the near grounding of *MS Viking Sky* (Figure 1) and the collision of *MS Dali* with the *Francis Scott Key Bridge* (Figure 2), reflect the inherent tension between the design choice of keeping the crew in control of the ship and the choice of protecting the ship's machinery from irreversible degradation or damage.



Figure 1. Viking Sky in a near-grounding vicinity to shallow banks [6, p. 98]. The auxiliary engines' automatic shutdown due to low lubrication oil pressure caused the ship to drift ashore, as it lacked power for its electrical propulsion.

Thus, unlike other industries with stationary assets (e.g., chemical processing plants), the use of alarms in the maritime context goes far beyond merely reducing demand on automatic layers of protection, in that for ships:

- 1 safety exists only as long as certain processes and functions are continuously controlled (e.g., power, propulsion, steering, and stability), and

- 2 that 'fail-safe' shutdowns of machinery merely propagate the situation to the crew, with the expectation of them recovering control.



Figure 2. MS Dali collision with the Francis Scott Key Bridge [7, p. 232]. The main engine shut down automatically due to low cooling-water pressure. As the main engine was no longer rotating the propeller, no wash was provided across the rudder. This meant that the crew could not effectively steer the vessel [7, p. 153].

This also means that *spurious trips* can constitute initiating events in and of themselves, as they too may lead to a degradation or loss of control.

4.2 Losing control of alarms

Over the last 3 decades, the maritime industry has seen a rapid rise in the number of strong 'attention-getting' alerts (alarms) provisioned on ships. We know this because senior seafaring officers report this phenomenon throughout their careers [1, pp. 90, 127]. While we cannot put exact figures against the seafarers' anecdotes, we can observe this growth by tallying required alarms and warnings (attention-getting alerts) in successive versions of the *IMO Code on Alerts and Indicators* (Figure 3).

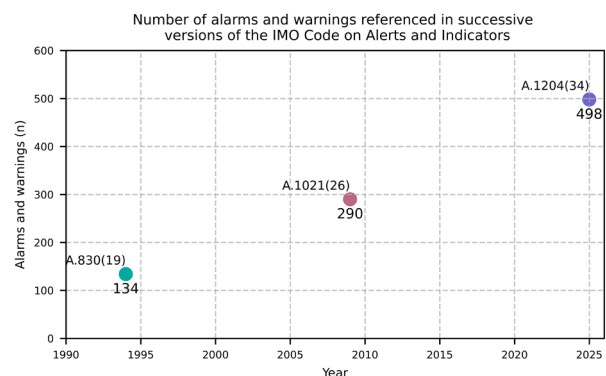


Figure 3. Number of alarms and warnings referenced in successive versions of the *IMO Code on Alerts and Indicators* (a collection of alerts from all instruments).

This increase has occurred despite IMO formally noting a *proliferation of alarms* as early as 1991 [8].

A similar rise in alarms was observed in the process and power industries several decades ago. The

⁵ In class rules, one can identify over +150 requirements for automatic protective or safety functions, indicating loss of availability (loss or degradation of control) without any prior alerting.

main reason was the introduction of digital control systems, which enabled many more alarms to be configured at very low cost. While seen at the time as a way to improve safety, the opposite has been observed [9].

Previous work shows that the number of configured alarms on ships reaches several thousand (on average, 32% of Tags in the ship's integrated automation system) [1, p. 56].

These alarms are configured among various pieces of equipment and plant, usually routed to central control positions such as the navigational bridge and the engine control room (ECR) (Figure 4).



Figure 4. Watchkeeping engineer in the ECR on a passenger ship carrying more than 8000 people.

At these centralised control positions, alarms are a common complaint among seafarers [1]. They frequently appear in incident reports, with design or operational shortcomings cited as contributing factors [10] [11] [12] [13, p. 8] [14, p. 14] [15] [16] [17, p. 18] [18, p. 42] [19, p. 39] [6, p. 6].

Thus, it is no surprise that the alarm system's behaviour impacts the crew's performance [1] [20] [21]. In normal operation, the 'overdosing' of alarms leads to desensitisation to their importance. In dire moments, alarm floods oversaturate the crew's cognitive and perceptual capacities (Figure 5).

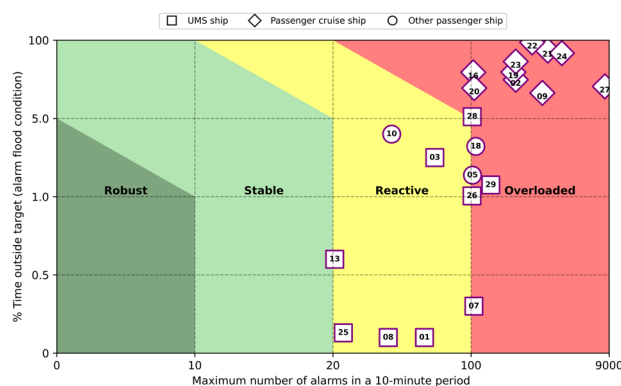


Figure 5. Alarm system performance against EEMUA 191 for upset conditions for participating ships in [21]. Courtesy of EEMUA⁶.

⁶ Note: ranges extended for the "Overload" category from 1000 to 9000 to be able to depict ship no. 27. And time outside target range extended beyond 5% to depict several of the participating ships.

5 APPROACH

The results presented in this paper are based on human performance data collected from actual events (incident investigation reports) and from simulator measurements in response to abnormal situations. Here, the operators' responses to design basis events were measured to assess the *timeliness* and *accuracy* of their actions. To minimise observers' influence on simulator participants, the data were automatically collected (e.g., positions of control elements from digital simulator logs and eye-tracking data) and later reduced using a combination of qualitative and statistical methods.

5.1 Incident data

In public incident reports, the factual information, such as timestamps of the crew's interventions and event logs, was analysed. In addition to the publicly available data, large shipping organisations kindly provided access to their internal incident databases and associated investigations (not publicly disclosed). Together, these records offered contextual insights into how the crew use their time responding to real-world abnormal situations. The incident data further revealed distributions of initiating event categories that lead to the abnormal situations (Table 1).

Table 1. Real abnormal situations analysed (blackouts, loss of propulsion, loss of manoeuvrability, etc.)

Initiating event category	Amount	Percentage
Spurious trip	81	43%
Mechanical failure	40	21%
Cyber event (software glitch/automation surprise, malfunctioning behaviour)	26	14%
Electrical failure	21	11%
External event (collision, grounding, shore power loss, fire)	12	6%
Maintenance activities ¹	7	4%

¹ Unintended event during maintenance due to omissions or errors of the maintainer

5.2 Naturalistic observations (field studies)

To better understand the time for crew to respond to alarms in operational environments, naturalistic observations were conducted on board three distinct large passenger ships.

5.2.1 Observations in the control room

Engineering officers participated in field studies while on watchkeeping duty. Here, they wore eye-tracking glasses (Figure 6).

Following the recordings, the watchkeeping teams shared their individual perspectives on the usefulness of each alarm and the causal and correlated relationships among the alarms encountered during the recorded watch.



Figure 6. An example of the eye-tracking recording material captured onboard a large passenger ship.

5.2.2 Observations of manual actions in response to simulated alarms

Recordings of simulated alarm responses in which the manual actions had to be performed outside the engine control room were also conducted for a range of alarms.

5.3 Simulator studies

Full-mission engine room simulators were used throughout this work:

- 1 K-Sim Engine MAN B&W 5L90MC VLCC L11-V: *Simulates the machinery setup of a VLCC.*
- 2 K-Sim Engine Diesel Electric Dual Fuel – CF II: *Simulates the machinery setup of a diesel-electric passenger ferry operating on MGO/HFO/LNG.*
- 3 Rheinmetall Electronics SES7: *Simulates the machinery setup of a large container ship.*

The simulators (Kongsberg and Rheinmetall) are all certified (e.g., to DNV-ST-0033) and are considered state-of-the-art in maritime simulation-based training (e.g., Figure 7).



Figure 7. The K-Sim Engine MAN B&W 5L90MC VLCC L11-V engine room simulator setup at the Copenhagen School of Marine Engineering and Technology Management (Teknika).

5.3.1 Design basis events (DBEs)

A collection of design basis events (DBEs; $n = 33$) was sampled across participating groups. The selection of DBEs was based on the simulator instructor's extensive seagoing experience as well as incident reports/data. Given the nature of dependencies among general (shared) services on ships, initial failures may propagate to other systems, escalating into a serious spiral of failures if not corrected in time (e.g., a loss of seawater cooling can trigger a cascade of failures across dependent systems such as propulsion and power generation).

Examples of end states (abnormal situations) include complete loss of electrical power (blackout), loss of propulsion (dead ship condition), flooding, explosion or fire. In several cases, the DBEs used in this study are collections of single points of failure or common-cause failure modes that have remained latent in ship design for decades.

5.3.2 Simulator measurements

In the simulators, studies used background scenarios representing a wide range of daily onboard work, from preparing engines and auxiliary systems for departure, bunkering, and underway transit to arrival at port. They also included stopping off machinery to conduct maintenance, inspection, or periodic testing. Throughout simulations, initiating events were injected by instructors from a separate simulator room, which simultaneously acted as the bridge of the ship (Figure 8).



Figure 8. The instructor's room acts as the bridge of the ship in the MC-90 simulator at Teknika. In the picture, the ECR is calling the bridge on the telephone. Instructors Jan Runge (right) and Martin Reimet (left).

One simulation study comprised several runs, each run lasting between 45 and 120 minutes, depending on the event and the group's aptitude. Participants wore eye-tracking glasses to assess their attention allocation. Simulator logs captured all manual actions on switch positions, control elements, and alerts.

After each run, a simulator file was also extracted, enabling a complete replay of the events, including all process variables, from a desktop version of the simulator. This allowed for post-simulation triangulation and analysis.

5.3.3 Simulator participants

Groups ($g=12$) of 3 to 4 persons ($p=39$) participated in the simulator studies. All participation was voluntary and based on written informed consent. Participants were not allowed to mix or to experience any design basis event more than once. This effectively meant that the groups were fixed. Due to time constraints and to gain breadth in the studies, groups were exposed to a variety of design basis events, which means that not all groups experienced the same alarms. This effectively prevented undesired transfer of learning between groups.

As it stands, the range of observed performance is more useful than the notion of an “average watchkeeper.” For that reason, participants were sampled from a broad range of skill levels, ranging from those undertaking their licensing exams to those with extensive maritime watchkeeping experience, several of whom held Chief Engineer’s Certificates of Competency, and endorsements covering electrotechnical functions (STCW III/2 and III/6, respectively). Not only did this establish a general understanding of how experience may influence a watchkeeper’s ability to respond to an alarm, but it also provides an initial reference point for the lower performance bound of the minimum competency level for inexperienced personnel. (e.g., “a newly qualified tradesman, but not an expert” in HEART [22]) who may shortly be placed in charge of the engineering watch. This is consistent with formerly used methods [23].

Simulations involving participants undertaking their licensing exams for engine room watchkeeping were sampled and conducted at the Copenhagen College of Marine Engineering and Technology Management and Fredericia College of Marine and Technical Engineering. The simulations with fully licensed operators were conducted at the Aarhus School of Marine and Technical Engineering.

5.4 Statistical reduction

Operator decision and response times were determined in post-analysis on the basis of the collected data. To do this, mixed methods were applied using:

- 1 *qualitative (subjective) inference by subject matter experts (simulator instructors with decades of seagoing experience as watchkeeping engineers)*

using supportive information like eye-tracking and simulator logs, as a basis for analysis, and

- 2 *objective (non-subjective) inference based on observable data alone, completion of a manual action within the critical sequence of actions (the “critical success path”) in response to the DBE.*

As discussed in later sections, either of these inference techniques carries its own limitations, including potential biases and uncertainties.

Using the annotated data, time intervals were calculated, and an empirical cumulative distribution function (ECDF) was developed to estimate confidence levels of crew responses (and behaviour) as a function of time. The models are intended to enable non-subjective evaluation of the cumulative probability of a successful diagnosis within a given available timeframe, as well as completion of time-critical actions within and outside the control room.

6 RELEVANT GOOD PRACTICE

This section explains the rationale behind the chosen approach and methods, particularly the decision to use both a superseded standard and its successor. Our retrospective analysis identified strengths and weaknesses across several standards and guidance related to taking credit for alarms. We discuss these findings and justify the selection of those deemed suitable for assessing transferability to the engine room context on ships.

6.1 IEC 61511 – Functional safety – Safety instrumented systems for the process industry sector

As a sector-specific interpretation of IEC 61508, IEC 61511-1 cl. 3.2.67 allows human action to be part of a safety instrumented function (SIF) within a safety instrumented system (SIS) [24]. Also, SIS safety requirements are required to include aspects related to the safety function’s response time:

- 1 “Response time requirements for each SIF to bring the process to a safe state within the process safety time.”⁷, and defines ‘process safety time’ as:
- 2 “Time period between the failure occurring in the process or the basic process control system (with the potential to give rise to a hazardous event) and the occurrence of the hazardous event if the SIF is not performed.”⁸

IEC 61511-2 A.8.2.1 [25] further informs:

- 1 “Any claims for risk reduction from an alarm layer of protection should be supported by a documented description of the necessary response for the alarm and justification that there is sufficient time for the operator to take corrective action, and assurance

⁷ IEC 61511-1 ed.2.1 “Copyright © 2017 IEC Geneva, Switzerland. www.iec.ch”

⁸ IEC 61511-1 ed.2.1 “Copyright © 2017 IEC Geneva, Switzerland. www.iec.ch”

that the operator will be trained (initially and recurring) to take the preventive actions⁹, and

- 2 “The system that undertakes the safety function would then comprise the sensor detecting the hazardous condition, the alarm presentation, the human response and the equipment used by the operator to terminate any abnormal situation.”¹⁰

6.2 EEMUA 191 – Alarm systems, a guide to design, management and procurement

EEMUA 191 3rd Edition [26] provides guidance on these 61511 requirements by presenting the reliability as a chain (Figure 9). An alarm functioning as a layer of protection is only as strong as its weakest link. For the ‘Operator Response’, emphasis is placed on the importance of proof-testing the alarm through scenario testing. However, there is no mention of any specific methodologies or references to good practices in designing and conducting such.

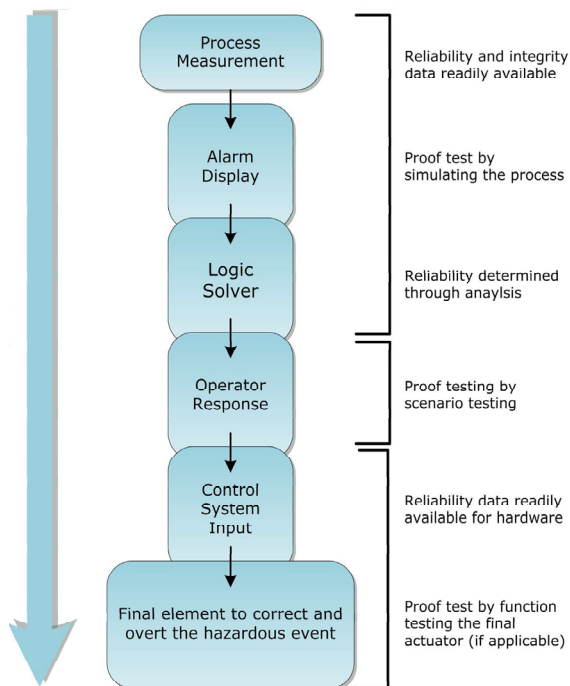


Figure 9. Proof testing of safety-related alarms against IEC 61511, EEMUA 191 3rd Edition, [26, p. 23 Fig. 10]. Courtesy of EEMUA.

6.3 IEC 62682 – The management of alarm systems for the processing industries

IEC 62682 provides some definitions of discrete time points and the intervals for operator response to alarms (Figure 10). These time intervals span from the initiating failure (i.e., a process variable like a temperature deviating beyond normal control envelopes) to the point at which operator action must be completed to prevent the consequence threshold from being exceeded.

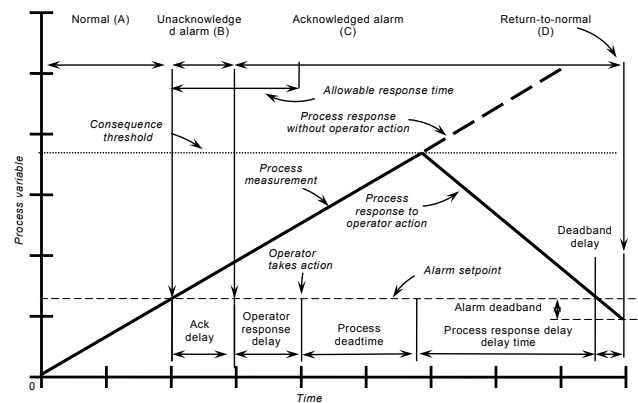


Figure 10. Alarm Response Timeline - IEC 62682:2022 [2, p. 33. Fig. 4]. Copyright © 2022 IEC Geneva, Switzerland, www.iec.ch.

As seen from Figure 10 several time delays, such as the acknowledgement delay, operator response delay, and process deadtime, are introduced, highlighting that the operator has less time to act than the *process safety time* defined in IEC 61511. With regards to the human-system interaction, the IEC 62682 5.4.6 explains the term ‘allowable response time’ as:

“The time beginning with the alarm annunciation and ending at the last moment operator action can prevent the consequence.”¹¹

Beyond this time interval, the consequence or exceedance of the design limiting requirement will occur even if action is taken.

Still, in Figure 10, it is apparent that some time delays are omitted. This includes omission of the indication delay (how long it takes the alarm system to detect and indicate the abnormal condition). Operator action is also shown as instantaneous, which is, of course, not the case in real life.

So, while IEC 62682 highlights the importance of considering ‘allowable response time’ for manual actions, it provides no specific methodology for estimating, measuring, or validating these time intervals.

However, various performance targets are provided, the majority of which relate to alarm system behaviour divided by time (e.g., alarms per hour). Here, one may get the impression that a control room operator can handle 6-12 abnormal situations (one every 6-12 minutes), as this is stipulated as a long-term performance target [2, p. 33. Table 7].

While we do not want to challenge the usefulness of the performance targets, as these have in fact

⁹ IEC 61511-2 ed.2.0 “Copyright © 2016 IEC Geneva, Switzerland. www.iec.ch”

¹⁰ IEC 61511-2 ed.2.0 “Copyright © 2016 IEC Geneva, Switzerland. www.iec.ch”

¹¹ IEC 62682 ed.2.0 “Copyright © 2022 IEC Geneva, Switzerland. www.iec.ch”

already proven useful in the maritime context [21], we deem them impractical for estimating the time people need to respond to alarm conditions (abnormal situations). This is because these targets fail to account for:

- 1 *the noise-to-signal ratio of the alarm system (people can deal with a virtually unlimited number of alarms if they are all false or irrelevant).*
- 2 *the degree of uniqueness between the alarms and the initiating event, some of which can trigger 100's of even 1000's of alarms (e.g., loss of electrical power)—yet the crew will still be responding to that single underlying event.*

We therefore deemed it necessary to base our methods on practices that consider operator response to the situation(s) generating the alarms, rather than the number of alarms themselves. Simply because uniqueness (only one alarm for one abnormal condition) is not guaranteed nor entirely possible to achieve with reasonable means.

6.4 NUREG CR 1278

The early standards development in the nuclear industry sought to address comparable issues arising from relying entirely on correct and timely human responses for specific design basis events.

This was identified by the nuclear industry as a generic safety issue at the regulatory level, needing a long-term resolution [27, pp. item B-17], and which had been previously raised as an action plan for safety research following the TMI-2 Accident, e.g., [28, pp. 78-79 (I.A.4)].

In 1983, the NUREG CR 1278 report (which became the seminal Handbook of Human Reliability Analysis) introduced the fundamental use of human performance times in risk assessments to account for human limitations, focusing on three core aspects:

- 1 *"How long it takes to begin a task (correct task initiation, including diagnosis of what to do),*
- 2 *How long it takes to do those tasks correctly,*
- 3 *The fixed time interval within which a task must be performed correctly."* - [29, pp. 200, ch. 6]

NUREG CR 1278 also discusses various sources and uncertainties in human response time data and lists a range of activities and lessons on the topic [29, pp. ch. 6, 8-2, 12, 14].

Some of those cited studies on operator response times later formed the basis of the ANS 58.8-1984 standard, which was later updated and adopted by

ANSI as ANSI/ANS 58.8-1994. Multiple frameworks for probabilistic human reliability techniques have since been developed, with only a minority seeing adoption and recognition by regulators [30].

6.5 ANSI/ANS 58.8 – Time response design criteria

We will briefly discuss key content and differences between the versions of the ANS 58.8 standards, including our rationale for applying a similar approach.

Either version is based on the premise that while operator models are generally recognised to be concurrent, iterative, and interwoven [31] [32] [33] [34] [35], time remains linear (at least in the Newtonian universe for which we care). So, in and of itself, the ANSI/ANS 58.8 approach applies a linearised sequence of discrete time points for human-system and system-process interactions, which generally occur during an unfolding DBE.

The period between these time points comprises time intervals that can be evaluated independently, facilitating analysis of response times for instrumentation and control elements, human elements such as diagnosis and manual safety actions, as well as the physical process itself.

6.5.1 ANSI/ANS 58.8-1994

The ANS 58.8-1994 version employs a calculation-based method to demonstrate performance, while also allowing field-based measurements as an alternative [23, p. 8].

Constants used for calculations are based on empirically collected data from simulator experiments and recorded events in which control room operators dealt with abnormal situations in real life.

Drawing on these empirical findings, the ANS 58.8-1994 working group established the standard to enable designers to assert a 95% confidence level for operator diagnosis within a minimum specified time (Table 2).

Table 2. Minimum $T_{\text{diagnosis}}$ for each plant condition¹².

Plant condition (PC)	Minimum diagnosis time
2	5 minutes
3	10 minutes
4 and 5	20 minutes

This minimum time is inferred based on the rarity or remoteness of the specific DBE and its resulting plant condition (Table 3).

¹² Extracted from American National Standard [ANS-58.8-1994 (withdrawn) (R2017) with permission of the publisher, the American Nuclear Society"

Table 3. Categorisation of events in terms of their likelihood of occurrence for the purpose of establishing safety criteria¹³.

Plant condition (PC)	Best estimate frequency of occurrence [F] per year
PC-1	Normal operations
PC-2	$F \geq 10^{-1}$
PC-3	$F \geq 10^{-2}$
PC-4	$F \geq 10^{-4}$
PC-5	$F \geq 10^{-6}$

A close analysis of the two tables reveals that conditions frequently anticipated (PC-2, Table 3) permit a diagnosis time as brief as 5 minutes (Table 2). In contrast, the situations leading to the rarest plant conditions (PC-5, Table 3) necessitate a minimum available diagnosis time of 20 minutes.

Rare events can pose considerable challenges for automation, as they are more likely to exhibit symptomatic cues and often require a highly contextualised (invariant) response. In such instances, the significance of safety claims for manual safety actions is markedly heightened. This is particularly relevant when the remoteness of the event enables designers to claim ALARP (As Low as Reasonably Practicable) with few controls.

Thus, it seems only reasonable to employ stricter requirements for the minimum available diagnosis time as designers categorise the event likelihood as increasingly rare: it is simply less likely that operators will have trained for or had experience with such events.

In the absence of actual performance times during early design phases, timing requirements are simplified to 1 minute for a discrete manipulation (Table 4). This is based on empirical data indicating 45-90 seconds when several discrete manipulations are required [23, p. 10].

Table 4. T_{operator} Sub-intervals (minutes) for each plant condition (PC) and for actions to be taken outside the control room. Categorisation of events in terms of their likelihood of occurrence for the purpose of establishing safety criteria¹⁴.

Plant condition (PC) ²	Fixed	Variable ¹
2	1 +	n
3	3 +	n
4 and 5	5 +	n
Actions outside the control room	30 +	n

¹ n signifies the number of discrete manipulations to complete a specific, single operator action.

² (Refer to Table 3 for plant conditions).

In addition to the manual actions, the ANS 58.8-1994 considers the time it takes for operators to make necessary preparations (acquire necessary tools, etc.) and reach the location at which the discrete safety action is to be performed.

Leaving the control room, consideration of the time required to put on protective clothing (e.g., a smokedivers suit) is likewise included, which is one reason for the 30-minute fixed interval [23, p. 5 cl.3.2.1/2]. Here, the core idea is to consider what the limiting case will be for the credited actions.

Lastly, the ANS 58.8-1994 requires designers to consider the delay times of the physical process, ensuring that systems do not begin to trip or exceed their design limits, even when all manual safety actions have been duly executed [23, p. 5 cl.3.1.2]. This recognises the time delays inherent in many thermodynamic processes, such as cooling, flow, valve stroke times, and so forth.

6.5.2 ANSI/ANS 58.8-2019(RE2024)

Instead of *safety-related operator actions*, the terminology and concept of *time-critical actions* (TCAs) are used. The ANS 58.8-2019 version applies a calculation-based method, prescribing a 50% time margin during design, as well as requiring simulation or field-based measurements (e.g. walkthroughs) for prelicensing demonstration of at least a 20% time margin [36, p. 9]. The performance evaluation formulas are simple and easy to use (Equation 1).

$$\text{Performance Time} = \text{(time from cue until action completed)}$$

$$\text{Available Time} = \text{(time from cue until analysis requires the action completed)}$$

$$\text{Margin} = \frac{(\text{Available Time}) - (\text{Performance Time})}{(\text{Available Time})} \cdot 100\% \quad 15$$

Our first concern in applying the ANS 58.8-2019 methods to ship design is that margins of 20% (or even 50%) can become increasingly insignificant (misleadingly optimistic) for DBEs, in which a short operator performance time is accompanied by a short time to consequence (exceeded design limit).

Say that for an arbitrary DBE, the total limiting time from the onset of the initiating event to the consequence threshold is 6 minutes (Table 5).

¹³ Extracted from American National Standard [ANS-58.8-1994 (withdrawn) (R2017)] with permission of the publisher, the American Nuclear Society"

¹⁴ Extracted from American National Standard [ANS-58.8-1994 (withdrawn) (R2017)] with permission of the publisher, the American Nuclear Society"

¹⁵ Extracted from American National Standard [ANS-58.8-2019 (R2024)] with permission of the publisher, the American Nuclear Society"

Table 5. A toy example of an overly optimistic TCA evaluation for a DBE requiring manual ESD activation.

Minutes from T-0	
0	DBE starts
1	Cue is given in control room ¹
2	A control room operator initiates ESD
(3!)	Time at which ESD must be initiated (point of no return)
3	ESD is completed
5	Process reaches a safe state
6	Limiting threshold if the process is not brought to a safe state

¹The cue may be an alarm or anything else that alerts the personnel that an event has commenced and that an action is required.

Say, the cue (alarm) triggers 1 minute after the DBE has commenced. Only a single operator action from the control room is required, e.g., manually initiating an ESD (Emergency Shutdown). This action is estimated to take at most 1 minute. Subsequently, the ESD sequence takes one minute to complete (e.g., stopping pumps, closing ESD valves), and the process variables will reach a safe state 1 minute after the ESD has completed.

In this example, the performance time is 1 minute (the time from cue to the required action being completed), while the available time (the time from cue until the manual action must be completed) is 2 minutes. While this yields a 50 per cent time margin, it leaves only 1 minute for detection and diagnosis (and plausible hesitation), which is, of course, unreasonably optimistic. Human reliability using human error probabilities (HEP) for diagnosing abnormal situations would estimate this as a 100% failure likelihood (Figure 11).

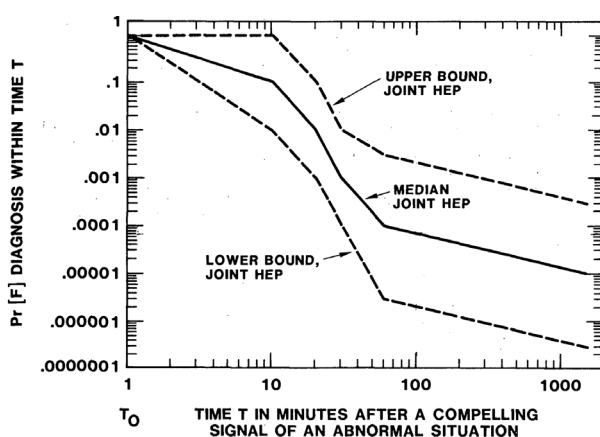


Figure 11. Nominal HEP of diagnosis time for abnormal situations (source: NUREG 1278) [29, pp. Figure 12-4].

The ANS 58.8-2019 version does not detail how to prevent the potential (mis)use error described in its proposed evaluation method. Instead, it (rightly) assumes that information displays (HMI), which alert and guide operators through DBEs and

associated controls, are designed according to established human factors principles (ANSI/IEEE 1023-2024(RE2019), and that the plant is designed to meet single-failure criteria (ANSI/ANS-58.9-2002(RE2015)) [36].

At present, no such assumptions concerning ship design can be made. And, despite some application of redundancy for certain systems on ships, the criteria for single failures are not consistently applied, and certainly not with consideration to operator errors or omissions.

Similarly, the *Center for Chemical Process Safety* (CCPS) quotes the *Upper Bound HEP* in Figure 11 (40 minutes) for the 99.9% probability of successful diagnosis, rather than the *Median HEP* (20 minutes) [37, p. 296]. We assume this is based on instructions when using the model, as it reflects the likely nature of the thousands of installed alarms in chemical process plants, for which bespoke, familiarisation-based training would never be practical. [29, pp. Table 12-5]

While the formula in ANS 58.8-2019 is straightforward, we were concerned that it still allows the design to proceed at risk if the 50% time margin cannot be established during the design phase, provided the designer is confident that the 20% time margin can still be verified and validated in practice [36, pp. 9 cl.5.2. item b, c, d]. This effectively means the asset is at risk of not being allowed to operate (i.e., obtain its license). As a result, measures need to be imposed to establish an adequate margin [36, p. 11]. If the 20% time margin subsequently still cannot be shown, the duty holder (asset owner) must justify why a lower margin is acceptable. That justification is then to be validated at least three times using different performers and more than one crew [36, p. 11].

In real-world maritime cases (sea trials, SRtP trials, and so forth) for which the 20% margin would not be satisfied, amending the design can entail significant costs and considerable delay in vessel delivery. Such will inevitably impose substantial commercial pressure upon those safety professionals responsible for the performance evaluation.

Although the ship's national administration and class societies are already delegated such levels of agency, we believe it would represent an organisational failure to consistently call upon field surveyors to exercise such subjective judgements regularly. For these reasons, our retrospective assessment led us to choose the ANSI/ANS 58.8-1994 approach of first establishing empirically valid response times in a maritime context to be used in design evaluations.

7 THE METHOD (DETAILED)

ANSI/ANS 58.8-1994(RE2017)'s Appendix 1, Section 6, outlines methods for data gathering and developing timing requirements in a similar manner that achieves consistent bases of methods.

For the simulator studies, this entailed selecting a set of DBEs and then evaluating the discrete time points, as shown in Figure 12, from left to right or right to left in the sequence of progression.

For the incident data, this consisted of marking up crew interventions described in the investigation reports or inferring them directly from the event logs within their supporting information.

7.1 Discrete time points

The applied discrete time points are shown in Figure 12 were defined as follows¹⁶:

- 4 **t_{St} : start of the event.** The time at which the DBE begins.
- 5 **t_{Ind} : Indication of the event.** The time at which information is readily available, e.g., one or more cues such as alarm signals and display indications alert operators that a DBE is happening.
- 6 **t_{Ack} : Acknowledgement of the event.** The observable time at which the indication of the DBE is acknowledged (usually logged in the alarm system). When acknowledgement is observed to be habitual (e.g., operator has a prior opinion that the indication is false), then t_{Ack} shall not be used in the analysis for the DBE under consideration.

- 7 **t_{ECA} : Earliest credited action.** The earliest time at which credit for the initiation of a time-critical action can be taken for any viable sequence (critical success path) of actions that will terminate the DBE or render it inconsequential for the safety, environmental or financial loss criteria of the ship.
- 8 **t_{MAI} : Manual action initiated.** The point in time for which analysis credits the initiation of a time-critical action.
- 9 **t_{SAC} : Safety action completed.** The time at which the time-critical action is evaluated to be completed.
- 10 **t_{SFC} : Safety-related function completed.** The time at which an indication is received that a safety-related system has performed its required safety-related function (a safety-related function can involve either automatic protection functions or manual operator actions or both).
- 11 **t_{Lim} : Event limit.** The earliest time at which a limiting design requirement would be exceeded if a safety-related function has not been completed.

7.2 Limiting design requirements

Design limiting requirements were defined as follows¹⁷: A limiting design requirement is the limiting value of a design parameter that ensures the consequence of any DBE does not result in:

1. "Unacceptable degradation ship of components, equipment, machinery, and

Definition of discrete time points and intervals

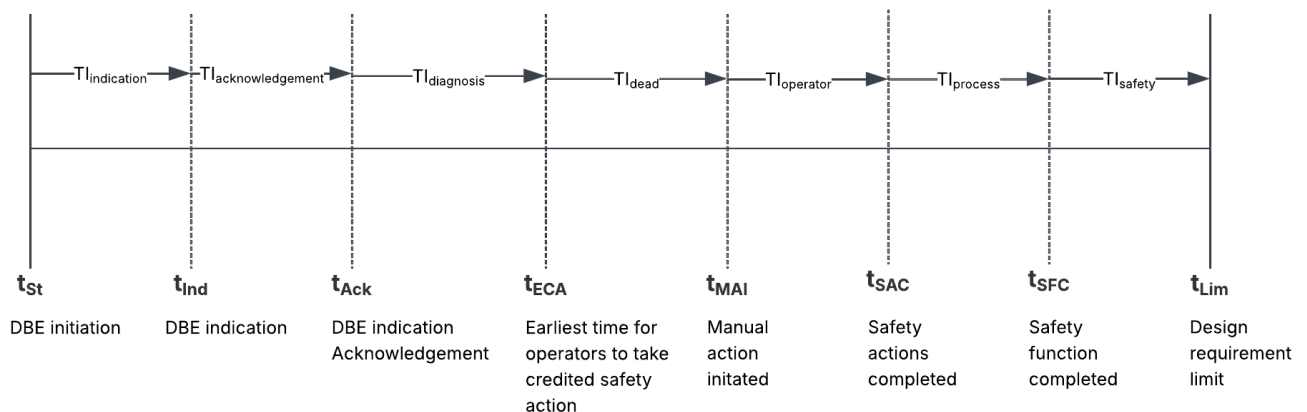


Figure 12. Discrete time points and intervals of a DBE progressing from onset (t_{St}) to the point of exceeding a limiting design requirement (t_{Lim}).

Note 1: Extracted from American National Standard [ANS-58.8-1994 (withdrawn) (R2017)] with permission of the publisher, the American Nuclear Society¹⁶

Note 2: Acknowledgement delay added as some ships operate with unattended machinery spaces (UMS)

¹⁶ Section extracted from American National Standard [ANS-58.8-1994 (withdrawn) (R2017)] with permission of the publisher, the American Nuclear Society¹⁶

¹⁷ Section extracted from American National Standard [ANS-58.8-1994 (withdrawn) (R2017)] with permission of the publisher, the American Nuclear Society¹⁷

systems required to mitigate or alleviate the consequences of any DBE."

2. *Loss of control over certain processes and functions for which safety exists only as long as control is continuous (e.g., power, propulsion, steering, and stability).*
3. *Violation of ship safety or environmental criteria as provisioned under applicable rules and regulations. Examples can include loss of life ranging from one to more fatalities, environmental damage, as well as loss of ship and cargo, and other damage (e.g., shore infrastructure).*

Once each time point has been marked, discrete time intervals (Figure 12) can be computed by taking the absolute differences between the discrete time points as defined:

7.3 Discrete time intervals

Elapsed time intervals between sequential discrete time points (Figure 12) were devised as follows¹⁸:

- 1 **$TI_{\text{indication}}$: indication delay.** Elapsed time between t_{St} and t_{Ind} .
- 2 **$TI_{\text{acknowledgement}}$: acknowledgement delay.** Elapsed time between t_{Ind} and t_{Ack} . The $TI_{\text{acknowledgement}}$ interval is to be disregarded if:
 - a. $TI_{\text{diagnosis}}$ starts immediately after $TI_{\text{indication}}$, since it can be assumed that the operator's decisions are already identified, or
 - b. Acknowledgement, t_{Ack} , is performed without further discrimination (e.g., the operator is observed to have a priori opinion of the alarm being false and delays response).
- 3 **$TI_{\text{diagnosis}}$: diagnosis time.** Elapsed time between system indication (t_{Ind}) (or the operator's acknowledgement (t_{Ack})) of the DBE and the earliest time for which credit can be taken for initiation of a time-critical action. During this interval, it is assumed that the operators recognise, discriminate and evaluate the DBE, verify automatic responses, observe system parameters, and devise subsequent actions. (Note: Acknowledgement is not the earliest credited action.)
- 4 **TI_{dead} : dead time.** The elapsed time between t_{ECA} and t_{MAI} , in which the analytic criteria permit operator action to be credited, but no time-

critical actions occur in the analysed sequences. During this interval, it is assumed that operators make necessary preparations (acquire necessary tools, etc.) and reach the location at which the discrete manual safety action is to be performed. A non-zero TI_{dead} may also occur following t_{SAC} (for any subsequent operator actions).

The first TI_{dead} interval may be eliminated if TI_{operator} starts immediately after $TI_{\text{diagnosis}}$ since it can be assumed that operator actions are already identified.

- 5 **TI_{operator} : operator response time.** The elapsed time between t_{MAI} and t_{SAC} . The time interval during which the operator initiates and completes time-critical actions.
- 6 **TI_{process} : process response time.** The elapsed time between t_{SAC} and t_{SFC} . The time interval between completion of a time-critical action and the indication that the corresponding safety-related function is completed through the response of the mitigating equipment, control elements, and the process itself.

For some DBE's, TI_{process} may be assumed to be zero (e.g., the opening of an air circuit breaker after pressing an emergency stop that de-energises its circuit).

TI_{safety} : safety time. The elapsed time between t_{SFC} and t_{Lim} . The time interval between completing the last safety-related function and when the event limit would have been reached without operator action and/or automatic action. A negative value of TI_{safety} is indicative of a limiting design requirement not being met (no time buffer).

7.4 Applications for multiple actions

The total sum of the time intervals for a sequence of events, each associated with a single time-critical action, is computed by adding the DBE's single $TI_{\text{acknowledgement}}$ and $TI_{\text{diagnosis}}$ to the TI_{dead} , TI_{operator} , and TI_{process} . In case of multiple operator actions, the multiple TI_{dead} , TI_{operator} , and TI_{process} are summed as¹⁹:

$$TI_{\text{Ackn}} + TI_{\text{diagnosis}} + \sum_{i=1}^n (TI_{\text{dead}} + TI_{\text{operator}} + TI_{\text{process}}) = \text{total time interval}$$

The computation of this total interval is necessary for each design limiting requirement identified for the DBE, with n being the number of operator

¹⁸ Section extracted from American National Standard [ANS-58.8-1994 (withdrawn) (R2017)] with permission of the publisher, the American Nuclear Society" 2 $TI_{\text{acknowledgement}}$ is added as some ships operate with unattended machinery spaces (UMS).

¹⁹ Section Extracted from American National Standard [ANS-58.8-1994 (withdrawn) (R2017)] with permission of the publisher, the American Nuclear Society"

actions necessary to prevent exceeding the limiting design requirements. For multiple operators, concurrence of $T_{I_{dead}}$, $T_{I_{operator}}$, and $T_{I_{process}}$ may be considered.

7.5 Examples of applied methodology

The method in ANS 58.8-1994(RE2017) can be complex to understand from its descriptions alone due to its formalised language. Examples are provided to show its use in the maritime context.

7.5.1 DBE simulator example (pre-trip)

A fundamental preventive DBE example is losing the seawater cooling function (primary cooling on a ship) due to a sudden restricted (dirty) sea chest suction filter (Figure 13). We present the annotated timestamps of group no. 4 (licensed watchkeeping engineers) responding to this DBE in Table 6.

Table 6. A DBE of seawater cooling loss used in the simulator studies (group no. 4, fully licensed).

time point	(min:sec)	description
t_{St}	00:00	Low suction sea chest filter 60% dirty
t_{Ind}	00:44	Alarm signal and indication for high diff. pressure over low suction sea chest filter
t_{Ack}	00:48	Acknowledgement of the alarm in the ECR
t_{ECA}	01:10	Earliest credited safety action – instructing a field colleague to open the high suction sea chest isolation valve
t_{MAI}	01:12	Starting to open the high suction sea chest isolation valve to re-establish cooling
t_{SAC}	01:19	High suction sea chest isolation valve opened
t_{SFC}	01:33	The time at which the re-established seawater cooling effect has reached end-consumers and halted the temperature rise
	12:38	The time at which the cooling flow must be re-established (point of no return)
t_{Lim}	12:52	The time at which the diesel generators will trip due to high temperature, causing a blackout and loss of propulsion (dead ship condition).

Using the annotated time stamps from Table 6, we can infer the time intervals (Table 7).

Table 7. Time intervals of the design basis event of cooling loss used in the simulator studies.

Time interval	(min:sec)	description
$T_{I_{indication}}$	00:44	Indication delay of 44 seconds
$T_{I_{acknowledgement}}$	00:04	Acknowledgement of high differential pressure alarm for the low suction sea chest filter takes 4 seconds
$T_{I_{diagnosis}}$	00:22 ¹	Diagnosis made 22 seconds after acknowledgement
$T_{I_{dead}}$	00:02	Operator moves to HMI location in 2 seconds.
$T_{I_{operator}}$	00:07	Operator takes 7 seconds to complete the opening of the valve on the HMI
$T_{I_{process}}$	00:14	Sea water cooling flow is re-established, and cooling reaches consumers in 14 seconds (temperatures stop rising and begin to drop shortly after)
$T_{I_{safety}}$	10:39	Remaining safety time was 10 minutes and 39 seconds

¹Readers should note that this is a simple DBE in the applied collection. A nearly identical DBE, where the same sea chest is obstructed by plastic or ice upstream of the filter (symptomatic as only cued by a cascade of high temperature alarms), yielded diagnosis times exceeding one hour for the same group (4208 seconds to be exact).

General to a centralised control context, and for this simulator example in particular, the $T_{I_{dead}}$ interval is negligible because the participants do not need to physically transition to the location of the isolation valve for the high suction sea chest filter.

Instead, this is done by taking a few steps to one of the HMI screens, where the action is accessible (Figure 14).

While manual tasks such as opening the seawater cooling isolation valve take a very short time in the simulator's environment, it is different in reality, as such valves may be large (Figure 15).

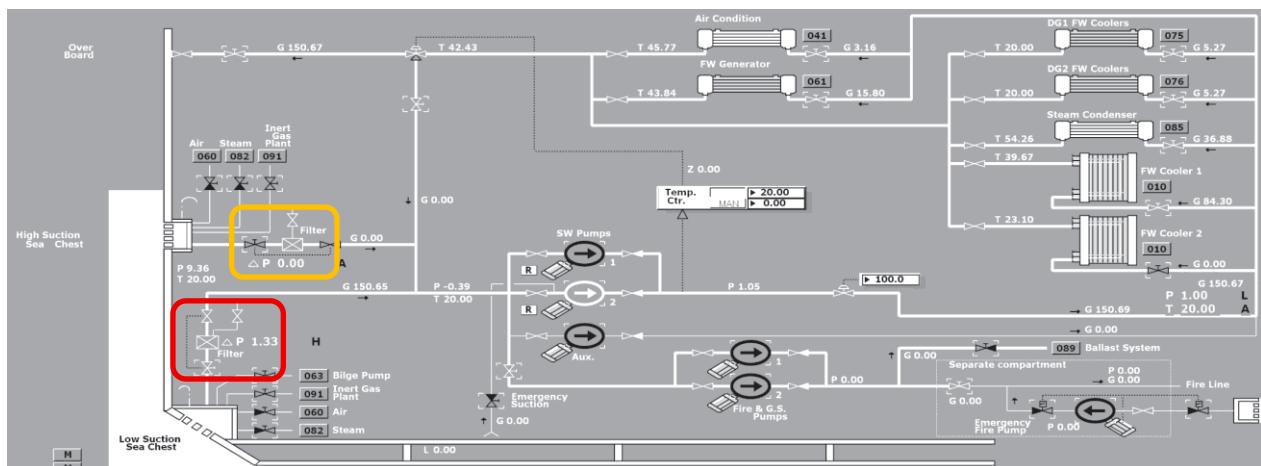


Figure 13. SCADA mimic of the seawater cooling system for the K-Sim engine room model of a VLCC. The DBE is initiated due to the blockage of the sea chest filter (red). The only manual safety action necessary is to bleed and open the isolation of the high suction sea chest filter (orange).

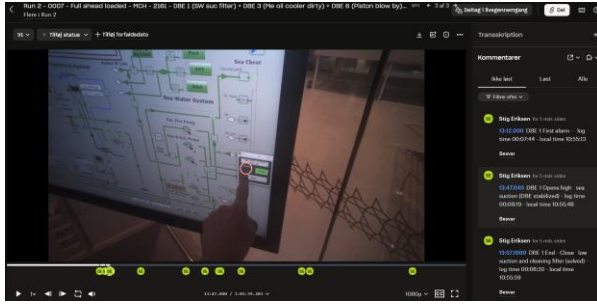


Figure 14. A participant presses the open button on the simulator HMI screen to open the manual high suction sea chest isolation valve (ref. Figure 13).



Figure 15. A sea chest filter with a manual inboard isolation valve on a passenger cruise ship. The valve outboard of the filter is remotely operated.

Although the simulators ‘near instant actions’ were considered a drawback at first, it turned out to be beneficial for analysing the $T_{diagnosis}$ interval. This is because this interval is completed on the initiation of the earliest credited safety action on any critical path of actions that restore seawater cooling flow. These actions were available from the simulator log files or from eye-tracking recordings.

7.5.2 DBE incident analysis example of a sequence of events (post-trip)

A notable real-world incident (post-trip) DBE is the MS Viking Sky, which experienced a blackout near the Norwegian coast due to insufficient lubrication oil levels in the sumps of its diesel generators, among other enabling conditions [6] [38, p. 26]. Given the rich detail in the formal incident report, we can approximate similar discrete time points from the timelines and narratives. Further, we highlight the application of analysing multiple sequential events (Table 12).



Figure 16. Point of t_{lim} (14:27:xx), (potential hull breach with subsequent flooding) [6, p. 27 Fig. 13].

Table 12. Low lubrication oil pressure trips on diesel generators on MS Viking Sky (enabling conditions: heavy weather; low lube oil levels in sumps)

time point	timestamp ¹ (hh:min:sec)	Comments
t_{St}	13:58:31	Blackout due to shutdown of DG1+2
t_{Ind}	13:58:41	Alarm flood (1000+ alarms) in ECR indicates blackout
t_{Ack}	xx:xx:xx ²	Not inferable
t_{ECA}	xx:xx:xx ²	(like very close to t_{MAI_1})
t_{MAI_1}	14:12:41	Transfer of lube oil from the storage tank to the diesel generators’ (primarily DG2) lube oil sumps commences ³
t_{SAC_1}	xx:xx:xx ²	Lube oil begins flowing to DG2 lube oil sump
t_{SFC_1}	xx:xx:xx ²	(Enough lube oil transferred to restart DG2 successfully without it tripping on low lube oil pressure)
t_{MAI_2}	xx:xx:xx ²	(local start of DG2 commences)
t_{SAC_2}	14:22:11	DG2 starts in local mode (30 sec.)
t_{SFC_2}	14:22:41	DG2 connected and feeding power to the main switchboard ⁴
t_{MAI_3}	xx:xx:xx ²	Restarting propulsion (propulsion motor) commences
t_{SAC_3}	xx:xx:xx ²	Propulsion motors restarted
t_{SFC_3}	14:37:xx	Propulsion motors are operational and provide enough thrust for the ship to move at 1 to 5 knots ahead ⁵
t_{Lim}	14:27:xx	The time at which the ship drifts over a rock at a depth of 10 metres at a draught of 6.5m and in seas with average wave heights between 4-6m (Figure 16)

¹Readers should note that the timestamps of the integrated automation systems can drift significantly over time, and hence, the use of such timestamps must be adequately calibrated, as was done by the NSIA [6, p. 16].

²Information not available or inferable.

³Inferred from the incident report narrative [6, pp. 25, 99].

⁴Inferred from the incident report narrative [6, p. 27].

⁵Inferred from the incident report narrative [6, pp. 25, 99].

In the absence of complete information, the *total time interval* may still be inferred as long as data is available as to when safety functions or essential services were completed/recovered. In the case of Viking Sky, this sums up to 39 minutes (Table 9).

Based on the description in the incident report, we infer a 14-minute diagnosis time of the engineering personnel onboard the Viking Sky (Table 13).

Table 13. Time intervals of the design basis event of the Viking Sky incident.

Time interval	(min:sec)	Comments
$T_{Indication}$	00:00	Indication starts with multiple cues (also an alarm flood)
$T_{Acknowledgement}$	xx:xx	Information not available
$T_{diagnosis}$	14:00	First credited action on the success path to mitigating the dead ship condition was to re-establish the lubrication function by increasing oil

		<i>levels in the generators' lube oil sumps</i>
$T_{I_{dead_1}}$	xx:xx	<i>Not inferable</i>
$T_{I_{operator_1}}$	xx:xx	<i>Not inferable</i>
$T_{I_{process_1}}$	xx:xx	<i>Not inferable</i>
...	...	...
$T_{I_{total}}$	39:00	<i>Operator and process response time including diagnosis</i>
$\sum_1^3 \left(T_{I_{dead_n}} + T_{I_{operator_n}} + T_{I_{process_n}} \right)$	25:00	<i>Operator and process response time excluding diagnosis</i>
$T_{I_{safety}}$	-10:00	<i>The last safety function of propulsion with sufficient thrust was to be re-established at least 10 minutes prior to t_{Lim}</i>

It highlights that the initiating event and enabling conditions (low oil sump levels) had to be addressed by the crew in order to terminate the subsequent events (loss of power and propulsion).

7.5.3 Field observation of DBE with actions outside the control room (pre-trip)

On board a large passenger ship, the eye-tracking recordings captured a DBE of a low-temperature freshwater cooling control valve sticking for one of the ship's diesel generators.

This caused a plethora of temperature alarms on the affected engine. Prior, the watchkeeping engineer had just stopped another generator due to problems with its exhaust gas boiler.

The observation serves as an example of the crew having a priori opinions regarding the relevance of an alarm and its impact on $T_{I_{diagnosis}}$ (time for correct task initiation).

Since the watchkeeper was accustomed to receiving a series of initial alarms upon starting the generator (due to transients), no action was taken at first (Table 10, Table 11, Figure 17).

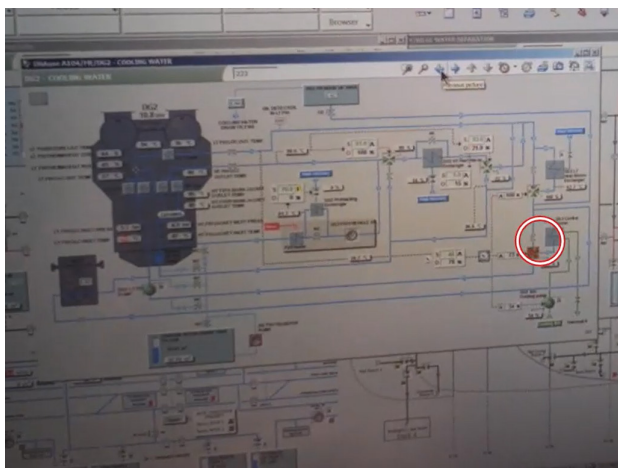


Figure 17. Watchkeeper eye gaze dwells on malfunctioning control valve for the DG LT FW cooling circuit.

Table 10. A DBE of a sticking control valve for the low temperature freshwater (LT FW) cooling circuit on a diesel generator.

time point	timestamp (hh:min:sec)	Comments
t_{St}	00:00:00	<i>DG2 started with a stuck LT FW control valve</i>
t_{Lim}	00:02:40 ¹	<i>The time at which the engine protection system will initiate load reduction, followed by an automatic shutdown</i>
t_{ind}	00:03:05	<i>DG2 Charge Air Inlet Temp High (cooled by LT FW Circuit)</i>
t_{Ack}	00:03:07 ²	<i>Watchkeeper acknowledges alarm without subsequent action (believing it to be an intermittent transient)</i>
	00:09:19	<i>Watchkeeper acknowledges high temp. alarm on LT FW circuit without subsequent action (also believing it to be an intermittent transient)</i>
	00:11:33	<i>Additional temp. alarms cues watchkeeper who now opens detailed mimic on DG2 live sensor readings</i>
	00:11:42	<i>Position error alarm is raised on the LT FW control valve</i>
	00:13:07	<i>Watchkeeper eye gaze dwells on faulty three-way LT FW control valve, then put it into manual (remote) control</i>
	00:13:28	<i>Watchkeeper informs 1st Eng. about issue with the control valve, then brings up detailed diagnostics mimic on the IAS HMI for the valve</i>
	00:15:40	<i>Watchkeeper and 1st Eng. discuss which HMI values are the control (setpoint) and which one is the feedback (position), as it's ambiguous</i>
t_{ECA}	00:15:49	<i>1st Eng. agrees to go down to the valve and exercise it physically through its range, so they can see which value is the feedback (position) on the HMI,</i>
	00:16:40 ²	<i>Watchkeeper receives phone call from 1st Eng. and confirms TAG-ID of the control valve, then hangs up</i>
	00:20:31	<i>Watchkeeper receives phone call again from 1st Eng. who is now ready to manually move the valve</i>
t_{MAI}	00:21:32	<i>The valve is manually exercised with the watchkeeper (ECR) vocally confirming feedback position values on the HMI with 1st Eng. over the phone</i>
t_{SAC}	00:23:05	<i>The control valve has been manually exercised through its full range</i>
t_{SFC}	00:23:13	<i>The LT FW temperature alarm return to normal</i>
t_{Lim}	00:23:34	<i>The valve discrepancy alarms return to normal, and watchkeeper ends the call³</i>

¹The crew operated with the main circuit breakers for each standby engine in manual mode and opened. This prevented usual intermittent alarms (triggering load reduction and shutdowns) from initiating the start-up of standby engines by the ship's power management system (PMS).

²Although acknowledged, the first indication was believed to be false.

³during the course of 23 minutes and 34 seconds, 103 alarms were raised in the ECR, of which 42 were related to DG2.

It also shows that users may override automatic protection systems (in a controlled manner) to keep the information available. In this example, the

temperature readings would become unavailable if the engine protection system had been able to take executive action and shut down the engine.

Table 11. Time intervals of the design basis event of the sticking control valve.

Time interval	(min:sec)	Comments
$T_{\text{indication}}$	03:05	Indication starts alarm on too high temperature for the LT FW circuit
$T_{\text{acknowledgement}}$	00:02	We discard t_{ack} in the analysis as discrimination is delayed (part of diagnosis)
$T_{\text{diagnosis}}$	12:44	Engineering team decides to manually (physically) exercise the control valve through its range of motion
T_{dead}	04:42	Time taken for 1 st Eng. to locate control valve, acquire tools and get in position to exercise control valve
T_{operator}	02:34	Time taken to manually exercise the control valve (in both directions) so that it can 'home-in' on its setpoint
T_{process}	00:08	Time taken for LT FW cooling temp. to reach normal operating envelope
T_{total}	23:34	Sensor, operator and process response time including diagnosis
T_{safety}	-20:33	The time at which the LT FW cooling temp. should have been normalised to avoid the engine protection system initiating automatic load reduction, followed by shutdown

8 RESULTS AND DISCUSSION

We structure the presentation of results according to the operator time intervals in Figure 12.

8.1 Acknowledgement time ($T_{\text{acknowledgement}}$)

In the simulators, we observed the time interval of $T_{\text{acknowledgement}}$ to be negligible, showing only a few seconds of delay from t_{ind} (Figure 18).

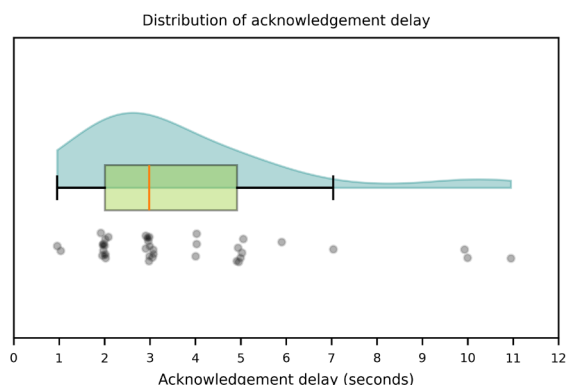


Figure 18. Distribution of $T_{\text{acknowledgement}}$ (simulator data).

With regards to ships operating with unattended machinery spaces (UMS), $T_{\text{acknowledgement}}$ will be another matter altogether, as the watchkeeping engineer on duty may be elsewhere than the ECR and, in some cases, resting in a cabin.

Conversations with crews on large cargo ships have revealed that the time it takes to reach the ECR from the cabins can be upwards of 7 minutes.

8.1.1 Impact of nuisance and false alarms

In field studies, we observed that watchkeepers do not initiate diagnosis upon acknowledgement (t_{ack}) if their mental model believes the alarm to be false. In such instances, they either sought or waited for confirmatory information. In one case, $T_{\text{diagnosis}}$ was extended up to 10 minutes (Table 11). Naturally, this behaviour is driven by the context of alarm systems' usability (or lack thereof), which also aligns well with established research on human factors [39].

8.1.2 User behaviours with several alarms active at once.

We observed that the user's gaze dwells first on the TagID (unique identifier) and alarm description (contextual information), and then the plant area. The watchkeeper would then navigate to a subscreen (screen in the screen) or adjacent HMI containing the process mimic of the alarm variable or system (Figure 22).

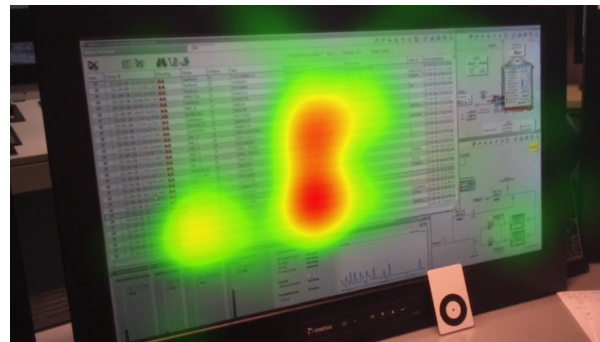


Figure 19. Dynamic heatmap of gaze fixations for the same alarm list HMI as above. Analysis shows heavy concentration on alarm messages and the area.

Given the high rate of alarms (82 to +100 per hour), it was unexpected that the watchkeeper's attention allocation (gaze fixations) seldom dwelt on the alarm's priority (Figure 19, Figure 20).

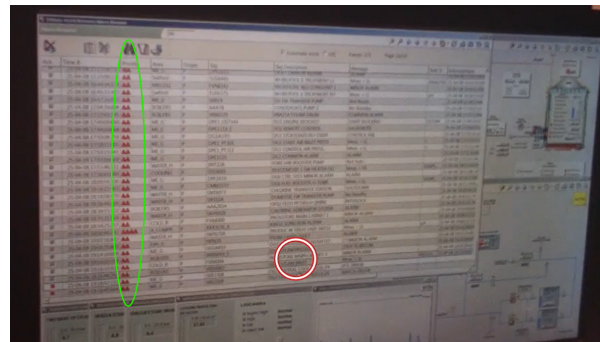


Figure 20. A watchkeeper's gaze dwells on the alarm list display during observational studies. The primary attention allocation (gaze fixations) was observed to be on the Tag description, the unique identifier (TagID) and the Area for the alarms. The priority column (green-circled) was given little to no attention.



Figure 22. Dynamic gaze trajectory mappings with fixation lengths between 300ms (smallest circles) and +1000ms (largest circles) on the alarm list HMI on a large passenger ship.

After all, the primary idea for alarm prioritisation is to help the user decide what to respond to first when there are too many alarms to deal with at once. Yet several high-priority alarms were not acted upon immediately, and several low-priority ones were. As we initially found those observations worth a deeper look, we conducted a time-consuming qualitative annotation of 15 consecutive alarms handled by the participants, selected from a randomly chosen time interval of the recordings (Figure 21).

These showed that watchkeepers did not dwell on the alarm's priority (0 dwell time for 37 of 51 instances). Whereas the opposite behaviour was observed for the Tag and Description (0 dwell time

in 6 of these 51 instances). When dwell fell on priority, it averaged 141ms vs 2277 ms for the Tag and its description string.

Watchkeepers explained that the alarm descriptors (Area, TagID, and Tag description and message), including the associated mimics, gave them more context on the alarm's importance and what to do.

Although a tad controversial, it appears that priority categorisation is less important to end-users than commonly thought. At least one cannot expect a higher-priority alert to be attended to first. It also appears decisive that alarm messages are understandable, as this is used for diagnosis.

Alarm list HMI attention allocation (Gaze dwell duration) qualitatively annotated

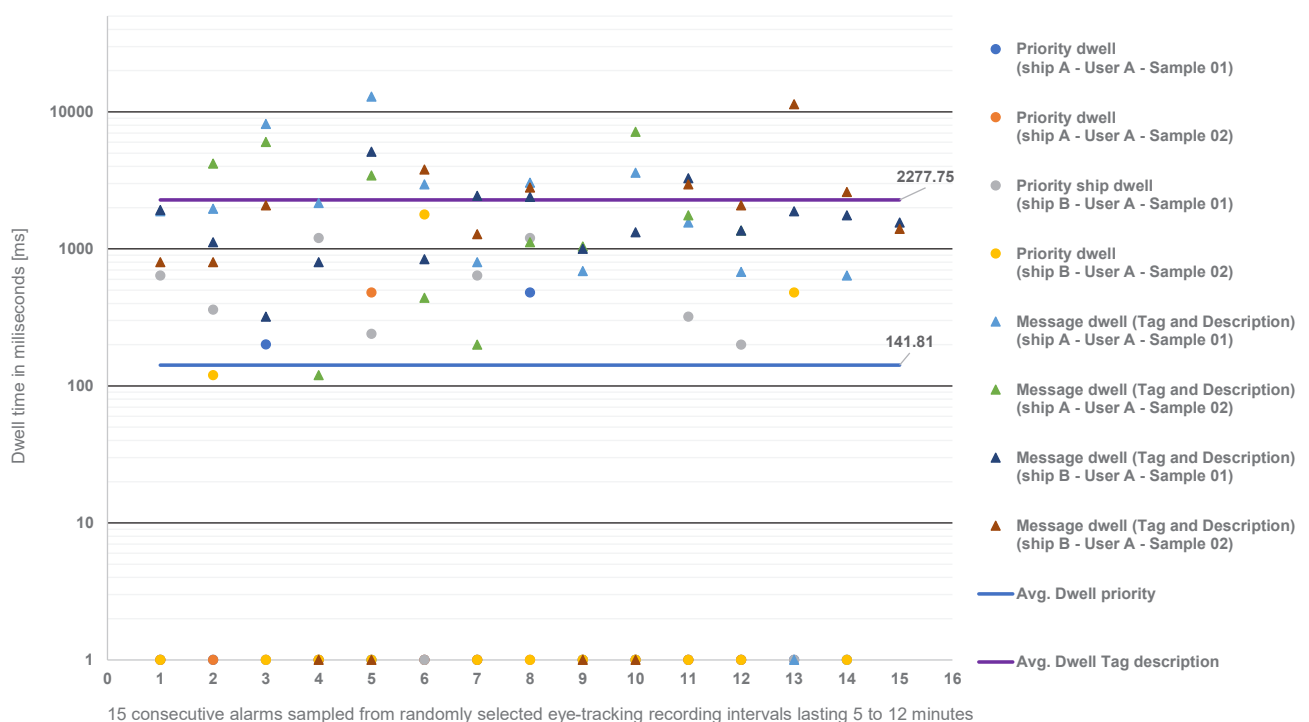


Figure 21. Alarm list HMI attention allocation (Gaze dwell duration) qualitatively annotated for 15 consecutive alarms observed in randomly sampled eye-tracking recording intervals lasting 5-12 minutes.

8.2 Diagnosis time ($T_{diagnosis}$)

A total of 187 real-world crew responses to incidents were analysed for $T_{diagnosis}$ intervals. Simulator data included 64 $T_{diagnosis}$ intervals from 12 participating groups across 33 design basis events for which no rationalised messages were presented.

Merging these intervals enabled the establishment of an empirical cumulative density function (ECDF), which was then transformed into a HEP model for sufficient diagnosis as a function of time (Figure 24). The HEP model parameters are specific to engineering personnel on ships. Use outside this context constitutes extrapolation and should be treated with caution. In contrast to the nominal diagnosis HEP within NUREG 1278, we've limited probabilities to a single decimal place (for maritime applications, it is irrelevant to speculate on human reliability to five decimal places).

We note that our resulting HEP diagnosis model largely falls within a factor of 2 of that presented in NUREG 1278, and that this resonates with at least one other study utilising simulators for 80 submariners undertaking their relicensing exams within the U.S. Navy [29, pp. 196 ch. 6-6]. On those grounds, our position is that the approach described in ANSI/ANS 58.8-1994's Appendix 1.6 [23] is suitable for developing and continuously refining minimum available timing requirements for time-critical actions by the engineering crew in

response to foreseeable initiating events within the ship's design basis.

8.2.1 Resonance between objective and subjective diagnosis times

In quantifying the resonance between objective inference of diagnosis time and expert judgment in the simulator data, we observed a tendency for experts to assign an earlier diagnosis than that inferred from observable actions alone (Figure 23). This is because simulator instructors can hear participants' discussions before the earliest safety action is initiated.

Discrepancy between objective and subjective inference of diagnosis time

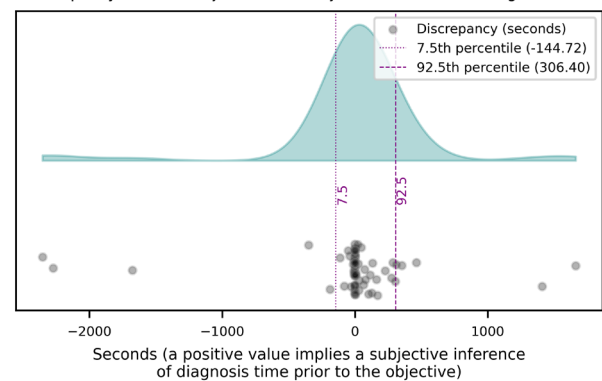


Figure 23. Difference between the objective and subjective estimation of diagnosis time by the instructors (simulator data).

In the opposite cases where subjective diagnosis lagged behind objective diagnosis, participants

Empirical model (1-ECDF) of diagnosis time by maritime watchkeeping engineering personnel following awareness of an abnormal event

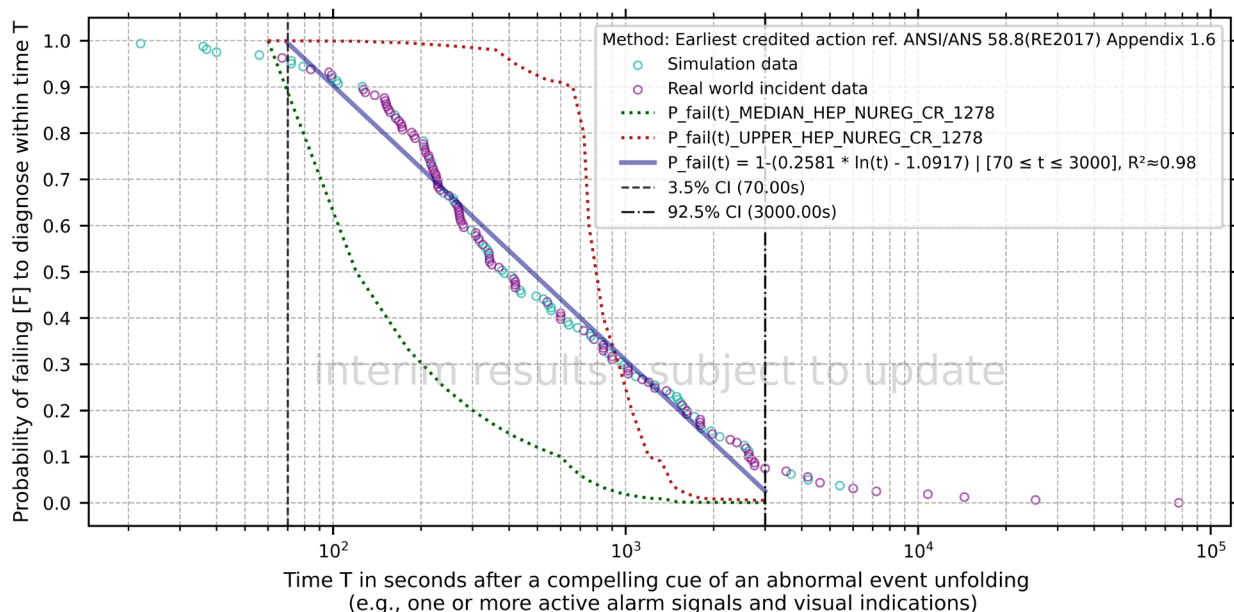


Figure 24. Model of diagnosis time based on the empirical data analysed in this study and using the earliest credited action as the inference time point of completed $T_{diagnosis}$. Data is based on non-rationalised alarm systems.

Note 1: The median and upper bound HEP presented in NUREG CR 1278 Figure 12-4 are depicted in Log-Normal form, to allow the empirically observed performance behaviour of marine engineering personnel to be interpreted in relation to the nominal median and upper HEP model used in the context of nuclear power plants (which are also that of current process safety practices [37]).

Note 2: This inclusion aims to prevent claims of equivalence when estimating maritime human performance times from industries where human factors engineering is considered on par with traditional engineering disciplines and is widely embedded in design and organisational practices.

initiated a credited safety action, hoping they were correct (or believing they lacked better options). Hereafter, participants typically observed the process's response to improve their understanding. In a few cases (~5%), a subjective diagnosis could not be made because participants failed to diagnose the situation, even though an initial corrective action was made (first lucky punch).

8.3 Operator errors and omissions

It's well understood that human beings make errors and omissions no matter how qualified they are. For the simulator studies, every single group made errors or omissions, some of which were recovered and some of which meant 'game over'. In the DBE of engine room flooding, quite some eyebrows were raised (Figure 25). During the simulations, we observed that:

- Groups quickly opened the emergency bilge suction (red, Figure 25). However, closing off the sea chest suction of the sea water cooling system often occurred with a delay or not at all (purple, Figure 25).

In several cases, it took time for the participants to realise that the hydrostatic pressure of the water column (draught around 16m) prevented suction through the emergency bilge valve (SDNR) (located on top of the tank top—at atmospheric pressure).

- Several groups omitted to manually override the controller's setpoint for the seawater cooling injection-circuit (orange, Figure 25) as well as opening the additional cooler for increased flow (blue, Figure 25).
- All groups that made it far enough into the DBE simulation forgot (omitted) to re-establish seawater cooling after having contained the

flooding. This caused a subsequent blackout due to high-temperature trips at the DGs.

- In the absence of a timely response to the DBE, a blackout is imminent (due to too high flooding levels causing short circuits in electrical cabinets). At this stage, only one group found out that the auxiliary pump (green, Figure 25) was supplied from the emergency switchboard. The remaining groups tried to isolate faulty circuits to get the generators and the general service pumps back online, with little success.

The engine room flooding DBE was not unique in offering numerous opportunities for errors or omissions. But its observation serves as an exemplary case as to why incorporating human factors into ship design could yield safer ships. It should therefore be carefully considered whether any credit should be given to alarms if a single operator omission or error will anyway result in the limiting design requirement being exceeded. The opposite entails claiming 100% reliability for a human.

8.4 Operator time intervals ($T_{I_{dead}} + T_{I_{operator}}$)

8.4.1 Fixed operator interval ($T_{I_{dead}}$) for actions inside the control room

We found that the fixed operator time interval ($T_{I_{dead}}$) in ANS 58.8-1994 (1 to 5 minutes) to be a reasonable estimate for the control room, as operators may be busy with other matters (paper work in particular), and may hesitate to act in cases where this results in financial liability or delays, or if decisions on the response is deemed to be taken at higher levels of seniority (e.g., the chief engineer, or designated person ashore).

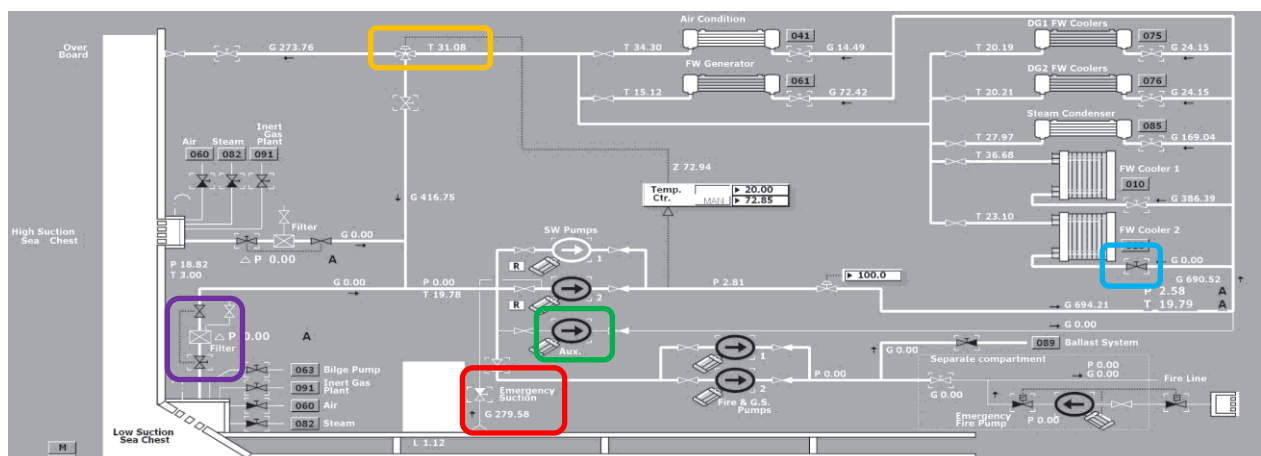


Figure 25. Engine room flooding DBE with marked-up manual action points. Since the seawater was cold (3 Celsius), the mixing valve controller recirculated about two-thirds of the cooling water to maintain 20 degrees Celsius at the coolers' inlets. This effectively means that only one third of the emergency bilge suction capacity was used.

Note 1: To avoid installing two separate pumps and separate piping, the emergency bilge system is integrated with the seawater cooling system.

8.4.2 Fixed operator interval (T_{dead}) for actions outside the control room

For actions required outside the control room, a fixed T_{dead} interval of 10 minutes seems reasonable, even for ships with permanently attended machinery spaces, where field personnel perform manual actions directed from the ECR. While it may be argued that field personnel are already near manual action locations, we observed that communication, access and coordination of manual actions require time. Often, far more time than the actual actions themselves. In some instances, accessing manual actions required physical keys. Keys which could only be collected in the ECR (Figure 26).



Figure 26. An example of a physical master key needed to access certain technical spaces on a passenger ship.

In terms of communication and coordination, language barriers and noise from running machinery often required multiple attempts to relay messages across the ECR. Sometimes, manual actions were delayed or hindered by day-to-day work. In one instance, access to manual valves was found obstructed by heavy objects placed on the floor plates above it (Figure 27).

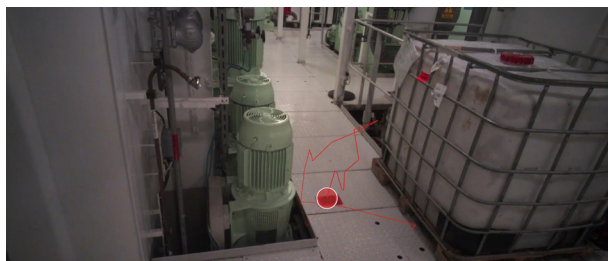


Figure 27. Access to the changeover valves for cooling the propulsion system was obstructed by a tank pallet.

In practice, it is not feasible to prevent such things with procedures. It's better that designers ensure such valves can be remotely controlled from the ECR, or that they are immediately accessible.



Figure 28. Engineer in the machinery spaces tracing the cooling system in search of a manual valve to be isolated as part of measuring the performance time for a simulated alarm response.

Finally, we observed that it can take several minutes to find valves located under floor plates, even for engineers with several years of experience on the given ship (Figure 28).

8.4.3 Variable operator interval ($T_{operator}$) for actions outside the control room

When field-tested on a sample of diverse ship types, the 1-minute constant specified in ANS 58.8-1994 for each discrete manipulation necessary to complete one safety action (e.g., activating a standby cooler takes four discrete manipulations), provides a fair estimate.

While it is a simplification, it does even out things not initially thought of in a task analysis (Figure 29).

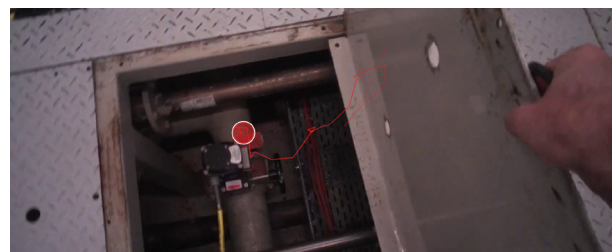


Figure 29. A manual safety action typically requires discrete steps that can easily be forgotten in analysis. The 1-minute constant evens out these uncertainties.

However, there were exceptions in which delayed maintenance or improper selection of materials had resulted in manual or remotely controlled elements being seized or hard to operate (Figure 30).



Figure 30. Example of a manual valve subjected to heavy corrosion by its environment and media.

It means that, for the 1-minute constant to hold, design considerations must be given to the materials and their operating environment, as well as to administrative programs such as preventive maintenance. The alternative is that these control elements might fail at times of need, as seen with the emergency bilge suction valve during the Emma Maersk flooding incident [12, p. Fig. 27].

8.4.4 Variable operator interval ($T_{operator}$) for actions inside the control room

For manual actions conducted in the control room, the 1-minute constant prescribed in ANS 58.8-1994

appears too conservative. On modern HMIs, discrete manipulations are swift (often a single mouse click). A quarter of a minute is more appropriate for mean time (Figure 31).

This is because watchkeepers can take time to search for the specific process mimic and the relevant control inputs. Likewise, they may also pause to verify that their manipulations were producing the desired effects before moving on to the next step in the response sequence (enaction).

8.5 Uncertainty and limitations of use

8.5.1 Estimating operator response times

A range of overly optimistic biases is well understood and documented for cases in which subjective opinions are sought for operator diagnosis time, as well as manual operator timings. Well-meaning and well-qualified individuals tend to underestimate the time required to perform manual safety actions by a factor between two and three [29, pp. 196, 235].

For diagnosis, it is likewise challenging for a design team to separate itself from the available information when determining the ease or difficulty of correct diagnosis and subsequent manual actions in real-world situations (hindsight is indeed one of the less valuable human virtues).

For these reasons, relying solely on subjective opinions regarding diagnosis and operator action timings carries significant uncertainty. As such, expert judgment applied in isolation is suitable only on an interim basis until actual data from the site or from simulations are available.

8.5.2 Using objective information to infer diagnosis time

The approach bases its inference of $T_{I_{diagnosis}}$ on the objective information available. This naturally carries a level of uncertainty, as it is not possible to read the minds of people. It also assumes that diagnosis happens prior to manual actions, when,

in reality, it is well known that operators gain more information as the systems and equipment under their control provide feedback throughout the unfolding event. Such feedback may constitute an integral part of the diagnostic process, especially for DBEs in which information for diagnosis is not available at the time of indication (t_{Ind}).

8.5.3 Using subjective opinion (expert judgement) to infer diagnosis time

For complex events requiring multiple operator actions, such as engine room flooding, it was at times challenging to define a clear dividing line between the end of $T_{I_{diagnosis}}$ and the beginning of $T_{I_{operator}}$. Notably, in cases where the earliest credited safety action was correct but the subsequent actions were delayed or omitted entirely (Figure 25), the data interpretation leads to large estimates of $T_{I_{operator}}$. For these reasons, expert judgment needed to be sought in a complementary manner in order to estimate the time point of complete diagnosis. Such qualitative analysis requires far more contextual data and analysis time for each interval of $T_{I_{diagnosis}}$.

Here, each case was backed by detailed notes referencing contextual information sources such as the eye tracking recordings (dwell bias, scan patterns, etc.), including speech transcripts (communication between the participants), simulator event logs, not to mention manual notes taken during each simulation from within the instructor room by an observer (a senior human factors specialist at LR).

8.5.4 Limitations

We understand that any criteria based on the time data presented in this study won't guarantee that the entire response will be correct. However, it is assumed that if there is adequate *available time*, then other performance-shaping factors (e.g., training, HMI layout, experience) will dominate the overall likelihood of failure to respond correctly in a

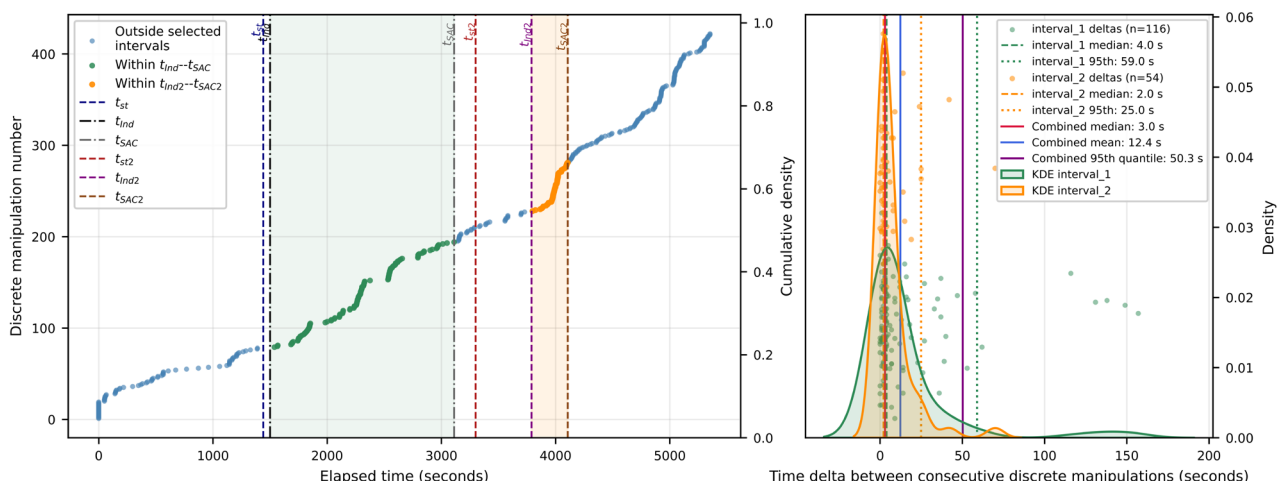


Figure 31. Example of time taken for the crew to complete discrete manipulations on HMIs in a centralised control station environment (full-mission simulator).

timely manner. Similarly, it is beyond the scope of this work to calculate actual operator response times. The data also do not define the limits of human capabilities, but rather document behaviour.

For this reason, it does not include the safety margins typically applied in system design and safety evaluations.

The factual information of response times from formal incident investigation reports of real events was incorporated into parts of this work, and no information was discovered that invalidates our findings in the simulators. The important point is that, when time models are developed from an accumulation of *actual* operator response times, operators, by design, should be able to respond in time.

9 CONCLUSION AND RECOMMENDATIONS

We set out to answer whether *seafarers have enough time to respond to alarms, and if not, what can be done to make sure they have going forward?*

We conclude that current general layouts and system designs on ships leave a lot to be desired in terms of allowing adequate time for the crew to carry out time-critical actions that fall within empirically valid response times. This is because:

- 1 *Diagnosis time extends up to 50 minutes merely to envelop 92.5% of the data points (Figure 24/Figure 11). This is attributed to a lack of understandable alarm messages and response procedures, which leaves the crew to their own guesswork as to what the problem is, what to do about it, and in what order.*
- 2 *Manual actions outside the control room are scattered and difficult to access. Performing them relies on long-term memory, as no response procedures detail their location or the necessary sequence (Figure 28).*
- 3 *Conventional designs provide little or no tolerance against errors or omissions which are foreseeable when relying on humans under stress (Figure 25).*

9.1 Going forward

Going forward, the industry must ensure that non-alarms²⁰ are removed, while ensuring that genuine alarms are justified and documented to meet the

criteria of being at *actionable, timely, reasonable, and trustworthy*²¹.

Industry recommendations

Based on the findings of this study, we advocate that all shipboard alerts be *rationalised* whenever these are intended to be presented to the ship's engineering crew. The rationalisation should take into consideration the alert's associated Design basis events (DBEs) in order to enforce *uniqueness*²². Since crediting alarms effectively means crediting manual actions, the rationalisation should be based on objective criteria that evaluate:

- 1 *Whether manual actions can be relied upon to initiate or control a specific safety-related function or must be accomplished by an automatic action,*
- 2 *when design modifications can obviate the need for manual or automatic actions, and*
- 3 *if credit taken for manual actions is reasonable in cases where a single omission or error results in exceeding a limiting requirement.*

During the design stage, the evaluation of high priority alerts can be performed using the time-response model in ANSI/ANS 58.8-1994 (Figure 12), in conjunction with time requirements derived from the methodology outlined in Appendix 1.6 of 58.8-1994, which yields Tables 12-14.

Table 12. Categorisation of events in terms of their likelihood of occurrence for the purpose of establishing safety criteria for time-critical actions (credited alarms).

Ship condition (SC)	Best estimate frequency of occurrence [F] per year
SC-1	Normal operations
SC-2	$F \geq 10^{-1}$
SC-3	$F \geq 10^{-2}$
SC-4	$F \geq 10^{-4}$
SC-5 ¹	$F \geq 10^{-5}$

¹ In relation to ANSI 58.8-1994, the SC-5 has been altered from 10^{-6} to 10^{-5} , as this is the target upper bound region for ALARP for new ships acc. to IMO FSA guidelines (MSC-MEP.2/Cir.12 Table 1).

Table 13. Minimum $T_{\text{diagnosis}}$ for each ship condition.

Ship condition (SC)	Minimum available diagnosis time ^{1, 2}
2	5 minutes
3	10 minutes
4 and 5	20 minutes

¹ The entire alarm system must be rationalised and incorporated under (audited) continuous monitoring and assessment.

² For probabilistic risk assessments, failure on demand for diagnosis is to be taken as: $P[F]_{\text{diagnosis}}(T) = 1 - (0.2581 \ln(T) - 1.0917)$, where T is the available time in seconds for the diagnosis in the limiting case for the DBE under consideration.

²⁰ Alert that does not indicate an abnormal condition which requires immediate (timely) action for the ships safety (see *IMO Code on Alerts and Indicators* for definition). A *time response design criteria* can reject non-alarms, as justification requires time-bound actions to be identified.

²¹ While we did not investigate criteria for *trustworthiness*, we emphasise its emerging importance based on our observations.

²² Abnormal situations in the real world rarely conform to single alarm simplifications, and our data reflect this reality.

Table 14. T_{operator} Sub-intervals (minutes) for each ship condition and for manual actions taken outside the control room (if necessary). Categorisation of design basis events in terms of their likelihood of occurrence for the purpose of establishing time criteria (see Table 12).

Ship condition (SC)	Fixed ¹	Variable ²
2	1 +	$0.25n$
3	3 +	$0.25n$
4 and 5	5 +	$0.5n$
Actions outside the control room ³	10 +	n

¹ Accounts for operator hesitancy as well as escalation of decision privileges (e.g., only senior officers will have the key to release CO2 during firefighting).

² 'n' signifies the number of discrete manipulations required to complete a specific time-critical action (e.g., putting a standby cooler into service usually requires opening 4 valves, which would entail 4 discrete manipulations).

³ If actions need to be completed in an unsafe environment (e.g., in a space subject to fire or toxicity), then add an estimate to equip PPE, as well as an estimate to egress to a safe area.

For preoperational validation, ANSI/ANS-58.8-2019(RE2024) provides an ideal basis for demonstrating adequate time margins, provided a fixed minimum time interval is added for an estimated diagnosis (see Table 13).

$$\text{Performance Time} = (\text{time from cue until action completed}) + T_{\text{diagnosis}}$$

$$\text{Available Time} = (\text{time from alarm indication (cue) until analysis requires the action completed})$$

$$\text{Margin} = \frac{(\text{Available Time}) - (\text{Performance Time})}{(\text{Available Time})} \cdot 100\%$$

It is advisable that performance validations take place before the ship enters a configuration where performing such actions could become necessary.

9.1.1 Recommendation boundaries

It is not our intention that other actions that have time limits that are not part of the ship's certification or licensing basis should be covered by these time criteria. Neither is it the intention that any such performance criteria supersede time criteria already included time-critical in preexisting regulatory instruments or classification rules. However, it may help justify changes to such instruments.

9.1.2 Not everything can or should be automated

We must further emphasise that it is not the intention to pursue automation for systems that exceed the current state of the art or become so complex that they introduce new safety risks without reasonable assurance that such automation of operator actions will have an overall beneficial effect on maritime safety. Sometimes

automation should not be authorised to respond automatically (e.g., CO2 fire extinguishing in the machinery spaces). In other instances, the correct (safe) response is difficult to determine and depends on a wide range of circumstances. An automatic system would quickly become very complicated and likely introduce additional risk. Here, a human operator will be better at evaluating the situation. In such cases, it may be more sensible to redirect efforts towards ensuring that both the alarm system and the operator's response are as optimal as possible, and/or by design modifications that obviate the need for automatic actions that would otherwise be required by adopting inherently safe design (e.g., using Halon instead of CO2 for firefighting in machinery spaces). In doing so, the available time may be reassessed to see if there is room for improvement, using sound engineering principles.

Examples include shortening the indication delay time by optimising instrumentation, minimising the number of manual actions needed, minimising the time it takes to reach local action positions, and optimising the physical process envelope to react faster to operator safety actions (e.g., better flow, better heat transfer, etc.), or extend the limiting design requirement by using tougher and more robust materials, and so forth. In either case, such design interventions can reduce required operator times, extend available time, or accomplish both. Such methods may involve existing design capabilities or adjustments to the design.

10 ACKNOWLEDGMENTS

Funding for this research was provided by the [Danish Maritime Fund](#), with substantial in-kind contributions from participating organisations.

11 DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in a live repository with runnable code and data at <https://deepnote.com/app/alarm-management/>.

Certain data supporting the findings of this study may be available on request from the corresponding author, [Asger Schliemann-Haug]. Data that are not publicly available are due to commercial confidentiality restrictions relating to incident data, and because the eye-tracking data contain sensitive biometric information from research participants that is subject to applicable data protection legislation as well as ethical restrictions.

12 WORKS CITED

- [1] Lloyds Register, "Lloyd's Register Alarm Management in the Maritime Industry - Volume 1 - A field investigation into the watchkeepers' experiences on watchfulness in a connected world," Lloyd's Register, London, 2024.
- [2] International Electrotechnical Commission (IEC), "IEC 62682 - Management of Alarm Systems for the Process Industries," European Committee for Electrotechnical Standardization, Brussels, 2022.
- [3] EEMUA - Engineering Equipment and Materials Users Association, EEMUA 191 - Alarm systems - A guide to design, management and procurement, 4th ed., London: EEMUA, 2024.
- [4] International Maritime Organization (IMO), "IMO Code on Alerts and Indicators," IMO, London, 2025.
- [5] International Maritime Organization (IMO), "DEVELOPMENT OF ENGINE CONTROL ROOM ALERT MANAGEMENT (ECRAM) PERFORMANCE STANDARDS," IMO Secretariat, London, 2026.
- [6] Norwegian Safety Investigation Authority (NSIA), "Report on loss of propulsion and near grounding of Viking Sky, Hustadvika, Norway 23 March 2019," NSIA, NO-2001 Lillestrøm, 2024.
- [7] National Transportation Safety Board (NTSB) (US), "Contact of Containership Dali with Francis Scott Key Bridge and Subsequent Bridge Collapse," NTSB, Washington D.C., 2025.
- [8] International Maritime Organization (IMO), "RESOLUTION A.686(17) - CODE ON ALARMS AND INDICATORS," IMO, London, 1991.
- [9] UK Health and Safety Executive, "Alarm Management," UK HSE, Liverpool.
- [10] Danish Maritime Accident Investigation Board, "Stena Scandica - Maine accident report on fire," DMAIB, DK-4220 Korsøer, 2023.
- [11] Marine Accident Investigation Board - UK, "Report on the investigation of the catastrophic failure of a capacitor in the aft harmonic filter room on board RMS Queen Mary 2 while approaching Barcelona," MAIB, Southampton, UK, 2011.
- [12] Danish Maritime Accident Investigation Board, "EMMA MÆRSK - Flooding of engine room on 1 February 2013," DMAIB, DK-2500 Valby, 2013.
- [13] Accident Investigation Board Norway (AIBN), "INTERIM REPORT 12 NOVEMBER 2019 ON THE INVESTIGATION INTO THE LOSS OF PROPULSION AND NEAR GROUNDING OF VIKING SKY, 23 MARCH 2019," AIBN, 2001 Lillestrøm, NO, 2019.
- [14] National Transportation Safety Board, "Collision between Cargo Ship Damgracht and AP Revelin," NTSB, 2023.
- [15] Maersk Line, "Maersk Eindhoven-103N encounters harsh weather Transpacific 6 service," Maersk, 19 February 2021. [Online]. Available: <https://www.maersk.com/news/articles/2021/02/18/maersk-eindhoven-103n-encounters-harsh-weather-transpacific-6-service>. [Accessed 23 February 2023].
- [16] Marine Accident Investigation Board (UK), "Collision between general cargo vessel Scot Explorer and gas carrier Happy Falcon," Marine Accident Investigation Branch (UK), 5 March 2024. [Online]. Available: <https://www.gov.uk/maib-reports/collision-between-general-cargo-vessel-scot-explorer-and-gas-carrier-happy-falcon>. [Accessed 14 March 2024].
- [17] Maritime Safety Investigation Unit (Malta), "Safety investigation into the grounding of the Maltese registered container ship Kea Trader," Maritime Safety Investigation Unit, MRS 1917, 2018.
- [18] Danish Maritime Accident Investigation Board (DMAIB), "UMOE VENTUS Fire on 23 December 2015," DMAIB, Copenhagen, 2016.
- [19] Australian Transport Safety Bureau (ASTB), "Engine room fire onboard MPV Everest - Southern Ocean on 5 april 2021," ASTB, Canberra, 2023.
- [20] Gauthier, Michelle & Kruithof, Gerard & Narlis, Christina & Jolliffe, Wendy., "Control and automation systems onboard the vessel: Lessons in human-centered design learned from 20 years of marine occurrences in Canada," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2019.
- [21] Lloyd's Register, "Alarm Management in the Maritime Industry - Volume 2 - An analysis of shipborne alarm data against the 62682:2014 performance metrics," Lloyd's Register, London, 2025.
- [22] J. Williams, "HEART - A proposed method for achieving high reliability in process operation by means of human factors engineering technology in Proceedings of a Symposium on the Achievement of Reliability in Operating Plant," Safety and Reliability Society (SaRS), NEC, Birmingham, 1985.
- [23] The American Nuclear Society and the American National Standards Institute, "time response design criteria for safety-related operator actions," The American Nuclear Society, Chicago, 1994(RE2017).
- [24] International Electrotechnical Commission, "Functional safety - Safety instrumented systems for the process industry sector," IEC, Geneva, 2016.
- [25] International Electrotechnical Commission, "IEC 61511-2 - Safety instrumented systems for the process industry sector - Part 2: Guidelines for the application of IEC 61511-1:2016," IEC, Geneva, 2022.
- [26] EEMUA - THE ENGINEERING EQUIPMENT AND MATERIALS USERS ASSOCIATION, Alarm systems - Guide to design, management and procurement, 3 ed., London: EEMUA, 2013.
- [27] U.S. NUCLEAR REGULATORY COMMISSION OFFICE OF NUCLEAR REGULATORY RESEARCH, "DRAFT DG-1052," U.S. NUCLEAR REGULATORY COMMISSION, WASHINGTON DC, 1996.

- [28] U.S. Nuclear Regulatory Commission, "NRC Action Plan Developed as a Result of the TMI-2 Accident," U.S. Nuclear Regulatory Commission, WASHINGTON DC, 1980.
- [29] A. D. Swain, H. E. Guttman, "Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications - Final Report," U.S. NUCLEAR REGULATORY COMMISSION, Albuquerque, 1983.
- [30] Health and Safety Laboratory (UK), "Review of human reliability assessment methods," Health and Safety Laboratory (UK), Derbyshire, 2009.
- [31] Wickens, C.D., "Engineering Psychology and Human Performance," Merrell, London, 1984.
- [32] Rasmussen, J., "Information processing and human-machine interaction: An approach to cognitive engineering," North-Holland, 1986.
- [33] Pew, R.W., Miller, D. C., Fehrer, C. E., "Evaluating nuclear control room improvements through analysis of critical operator decisions," in *Proc. 25th Annual meeting of the Human Factors Society*, Santa Monica, 1981.
- [34] Hollywell, P., Electrowatt Engineering Services UK Ltd.; Marshall, E., Synergy, "An experiment to support the design of VDU-based alarm lists for power plant operators," in *Proc. Ergonomics Society conference, The Human Factors of Alarm Design*, Birmingham, 1992.
- [35] Stanton, N., Baber, C., "Alarm initiated activities: an analysis of alarm handling by operators using text-based alarm systems in supervisory control systems," *Ergonomics*, vol. Vol. 38, no. No. 11, pp. 2414-2431, 1995.
- [36] The American Nuclear Society and the American National Standards Institute, *Time Response Criteria for Manual Actions at Nuclear Power Plants*, 2019 ed., Chicago, Illinois: American Nuclear Society, 2019.
- [37] American Institute of Chemical Engineers - Centre for Chemical Process Safety (CCPS), "INITIATING EVENTS AND INDEPENDENT PROTECTION LAYERS IN LOPA," John Wiley & Sons, New York, 2015.
- [38] Lloyds Register, "Viking Sky Investigation - Appendix G to the NSIA final investigation report," Lloyds Register, Southampton, 2023.
- [39] J. H. Kerstholt, "Decision making in a dynamic situation: The effects of false alarms and time pressure," *Journal of Behavioural Decision Making*, vol. 8, no. 3, pp. 180-200, 1995.
- [40] Olivia Walker, Lloyd's Register, "The human element competency required for design appraisal," Lloyd's Register, London, 2011.
- [41] International Maritime Organization (IMO), "Seafarers Training, Certification and Watchkeeping (STCW) Code - Annex 1 - Part A - Mandatory standards regarding provisions of the annex to the STCW Convention - Chapter III - Standards regarding engine department," in *STCW Code - Seafarers' Training, Certification and Watchkeeping*, London, IMO, 1995.
- [42] UK Health and Safety Executive, "Appendix A - Process for Managing Instrumented Alarm Functions credited with risk reduction against MAH's," UK HSE, Liverpool, 2018.