Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio

doi: 10.3850/978-981-18-8071-1_P182-cd

Cognitive Workload when Novices and Experts Supervise Autonomous Ships – Findings from Empirical Studies

Magnhild Kaarstad, Alf Ove Braseth
Institute for Energy Technology, Norway. E-mail: magnhild.kaarstad@ife.no; alf.ove.braseth@ife.no

Gudbrand Strømmen, Jon Bernhard Høstmark
Kongsberg Group, Kongsberg Maritime, Norway. E-mail: Gudbrand.Strømmen@km.kongsberg.com: Jon.Bernhard.Hostmark@km.kongsberg.com

In the maritime industry, there is currently a drive towards more environmentally friendly operations and reduced costs while maintaining a high level of safety. It is expected that the next major change in this domain will be autonomous or partly autonomous ships supervised from a land-based operation center. It is therefore a need to develop a safe and efficient operational concept. One research topic is to investigate how operator workload is affected in different situations. In this paper, we explore the following research questions: 1) How is workload experienced when supervising one vs three autonomous ships? 2) How is workload experienced by novices (gamers) and experts (navigators) while supervising autonomous ships? 3) How is workload experienced by experts while supervising three autonomous ships in different interaction design solutions? The questions are explored through empirical studies. Two maritime simulation exercises with novices and experts as participants were conducted. The findings indicate that workload is higher when supervising three ships compared to one ship. The findings also suggest that different display design concepts affect navigators´ situation understanding, and that some interaction design solutions are particularly challenging for novices. Findings from the study can be used to further guide interaction design development for supervising autonomous ships, and as a first step to explore competencies needed by future navigators.

Keywords: Autonomous ships, Remote operation center, Workload, Experts vs. Novices, Interaction design layouts, Situation understanding, Empirical study.

1 Introduction

In the maritime industry there is currently a drive towards more environmentally friendly operations and reduced costs while maintaining a high level of safety (e.g., Porathe et al., 2018; Kaarstad & Braseth, 2020). At the same time, the past decades have shown an increased investment in new technologies and automation (ibid.). A consequence of this is the on-going development of autonomous and semi-autonomous ships that can sail with less fuel consumption, an estimated reduction in operating costs, as well as increased safety. With this development, Remote Operation Centers (ROCs) are expected to be established, from where these ships can be monitored, and controlled if necessary. In the maritime domain, the international maritime organization (IMO, 2018) has proposed four levels of autonomy. The first level represents ships with automated processes and decision supports, and in level two, the ships are remotely controlled. In these two levels it is expected that seafarers are present in the vessels to take manual control if necessary. Level three represents remotely controlled ships without seafarers on board. In level four, the ships are fully autonomous, and the operating system of the ship is expected to make decisions and determine actions by itself.

When introducing autonomy and remote supervised vessels, the work tasks of the navigators will change considerably from today´s situation. One change will be the possibility of supervising more than one ship simultaneously. There is, however, a need to maintain a high level of safety and efficiency and a manageable workload for the operators in the ROC. It is therefore a need to perform research into intuitive and well-designed information displays that can be implemented in the operational concept for a ROC.

One such initiative is a project supported by the Research council of Norway (RCN) through "Land-based Operation of Autonomous Ships" (LOAS). The main objective of this project is to iteratively develop and test interaction solutions for a ROC ensuring safe and efficient supervision of autonomous ships. The project is focusing on identifying what information is needed in a land-based control center to remotely supervise autonomous ships, and how this information should be presented to support workload and situation understanding of future operators in a ROC. This paper is a contribution in the LOAS project.

The topic in this paper is how workload is experienced in different situations using different information display concepts. We ask the following research questions:

1) How is workload experienced when supervising one vs three autonomous ships?
2) How is workload experienced by novices (gamers) and experts (navigators) while supervising autonomous ships?
3) How is workload experienced by experts while supervising three autonomous ships in different interaction design solutions?

In the next chapter we present relevant background on the concept of workload and automation in relation to operational settings and competence level. Chapter 3 explains the user study, while chapter 4 and 5 present and discuss the findings. Lastly, conclusions and topics for further work are outlined.

2 Workload and Supervision of Autonomous Ships

One classic definition of mental workload by Hart and Staveland (1988) is "the perceived relationship between the amount of mental processing capability or resources and the amount required by the task". For the ROC operational concept, we need to develop knowledge related to how cognitive workload is perceived in different situations while supervising autonomous ships.

2.1 Workload and multi-unit supervision

Multi-unit supervision means that a team or an operator can monitor and control two or more units at the same time. There is currently limited experience of how workload is affected while supervising multiple units in general, and multiple autonomous ships in particular. Maintaining safety should be the goal when supervising ships remotely. Størkesen (2021) suggests that safety is not properly managed before safety management requirements and core tasks are aligned. In remote operations, there may arise new safety management challenges due to fragmented organizational structures, new role inter-dependencies and high workload on remote supervisors (ibid.). Wróbel, Gil, and Chae (2021) performed an expert-based study and found that inappropriate planning and inadequate supervision is regarded as less safety-threatening in remote operation than failure to correct known problems. A possible explanation to this finding was that failures in planning and supervision can be mitigated during the actual operations of remotely controlled vessels, and if the work is organized and distributed reasonable between the different persons in a ROC, there is a possibility to plan for an appropriate level of workload to avoid challenging situations (ibid.).

An important issue to consider related to multi-unit supervision, is that vessels in different modes (e.g., transit, berthing, unberthing) may impact how many vessels the operator can handle simultaneously.

In the aviation domain, sectors are combined, separated, and managed depending on traffic volume/complexity and the air traffic controllers' workload (Zelinski & Lai, 2011). In quiet periods, staffing is reduced, sectors combined, and each air traffic controller is responsible for a larger part of the airspace. In more hectic periods, the airspace is split into smaller sectors so that each air traffic controller has a limited and manageable area of responsibility. This solution requires predictable variations in the workload so that the total staffing in the control center is always adapted to the need (Eitrheim et al., 2019). Insight into the workload situation for managing more than one ship is required for the ROC operational concept.

2.2 Workload and expertise

The International Convention on Standards of Training, Certification and Watchkeeping for Seafarers (STCW, IMO, 1996) has established a set of competences for masters, officers, and watch personnel of maritime vessels. It has been noted that introduction of autonomous ships will lead to new work processes for navigators, and that it will be necessary to revise and adjust the

STCW competence requirements to be applicable for operators supervising autonomous ships from a ROC (e.g., Reling, et al., 2018).

Navigators have many years of practical maritime experience. According to Ericsson et al. (2006), expertise refers to the characteristics, skills and knowledge that distinguish experts from novices and less experienced people. Novices tend to have limited knowledge and lack the experience to recognize patterns, while experts have a deep understanding of their field and can recognize meaningful patterns of information quickly. Novices are often overwhelmed by the complexity of a problem, while experts are able to flexibly retrieve important aspects of their knowledge with little attentional effort (Chi et al., 1988).

When autonomous ships are supervised from a remote operation center, there is a risk these practical experiences gradually will be lost. In “Ironies of automation”, Bainbridge (1983) suggests that the more we automate, and the more sophisticated we make that automation, the more we become dependent on a highly skilled human operator. She argues that in a highly automated system, there are two roles that are left for the humans: to monitor that the automated system operates correctly, and to take over control if it fails or behave unsafe. Unfortunately, human skills deteriorate when they are not used. This means that a former experienced navigator monitoring an automated process, over time can experience a decline in skills. In addition, the situations when the operator will be called upon will by their nature often be cognitive demanding (Bainbridge, 1983). Veitch and Alsos (2022) performed a systematic review of 42 studies on human supervision and control of autonomous ships and found that human operators have a similar active role in ensuring safety in autonomous ship as in conventional ships. The ROC operators will need to take on a safety role above and beyond a backup role (ibid.).

For efficient and safe operation by operators in the ROC, the competence level of the ROC operators needs careful attention, and the information content and presentation must be suitable for the operators` competence level.

2.3 Workload and interaction design

One important way of supporting operator understanding is by developing intuitive interaction design displays. Several technological systems have been developed to support navigators, but these systems have not necessarily resulted in lower operator workload. Lützhöft (2004) found that a container vessel produced in the 1960s, and converted into a passenger ship in 1990, had 15 different suppliers of bridge systems, and an offshore supply ship built in 2005 had close to 30 different suppliers (Lützhöft & Nyce, 2014). Such designs will affect the operators´ workload as they need to integrate information from several different sources to operate in a safe manner. Several maritime accidents have been directly linked to mental workload of the seafarers (e.g., Hetherington et al., 2006). It is therefore a need to consider how workload is best supported in the interaction design solutions for the ROC operational concept.

To support a relevant level of cognitive workload, situational understanding is essential. Endsley (2013) describes situational awareness (SA) as consisting of three levels. SA1 concerns perception of the elements in the environment, SA2 is how the situation is comprehended or understood, while SA3 is the projection of the situation into near future status. (Endsley, 2013). This means that those with high situational awareness have not only perceived relevant information in their surroundings but are also able to integrate the information to understand its meaning and are able to project possible future scenarios based on this information (ibid.). A design principle in the LOAS project is to support all three levels of situational awareness, so that operators who are to monitor autonomous ships can understand the maritime picture with the lowest possible cognitive load.

3 Method

3.1 Participants

The findings in this paper are based on two empirical studies. In study 1, there were 12 participants. Three worked as navigators at a passenger ferry, five worked as vessel traffic center (VTS) operators. The last four participants were novices and students in their last year of maritime engineering. The novices were invited to provide user feedback on different interaction design layouts and offer suggestions for ways to make the system intuitive and user-friendly. The group of novices included in our study were advanced gamers. Although the gamers possess a range of cognitive abilities, they have little or no maritime experience. In study 2, there were eight participants. Four worked as navigators at a passenger ferry, and four worked as VTS operators.

The VTS operators in both studies had a former background as navigators. Therefore, the navigators and VTS operators in both studies are considered as experts. The participants were recruited through a contact-person at each location (the ferry company, the vessel traffic center, and the university) who were asked to provide the study with volunteers. All participants had completed, or were in the final phase of completing, a degree equivalent to a bachelor's degree.

The average age of the experts in study 1 was 50,5 years while the average age of the novices was 26 years. The average age of the experts in study 2, were 42,6 years. The average maritime experience of the navigators in study 1 were 30 years (this experience ranged from 14 to 50 years), while in study 2 the average maritime experience was 22 years (with a range from 8 to 40 years).

3.2 Scenarios

Two scenarios were designed for the studies. The scenarios were videotaped in advance. One scenario was developed for supervising one autonomous ship, and one scenario was developed for supervising three autonomous ships. The scenarios were designed to mirror a normal crossing of the Oslo-fjord, with realistic traffic and complexity. The videotaped outlook from the bridge represented a light foggy day, which encouraged the participants to pay attention to the digital displays available (e.g., radar, electronic map). In study 1, two different design layouts were used for the one-ship condition (A and B) while three different design layouts were used for the three autonomous ships condition (C, D and E). In study 2, the scenario from study 1 for the three-ship condition was used, and two different design layouts were compared (E from study 1 and a new design, F), see Figure 1. Each scenario lasted for about 20 minutes. The order of the scenarios was randomized within each group.

3.3 The interaction design layouts

The different interaction design layouts are presented in Figure 1.

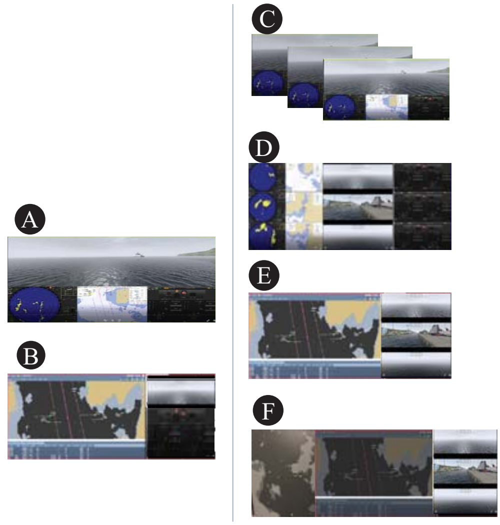

Figure 1: The different interaction design layouts (A, B, C, D and E was introduced in study 1, while E and F was tested in study 2).

In “A” the information is displayed as a “out of the window”-view, combined with separate displays for radar and electronic map at the bottom of the screen. In “B”, a display with radar information integrated with the electronic map is the main display, with a small “out-of-the-window-view” to the right. “C” is a three-ship setting, where the setup from A is used, and the information is flipped between the three autonomous vessels in 20 seconds´ intervals. In “D” the setup from A is used and the three ships are displayed in a parallel format. “E” shows three ships in a setup similar as in B, having one “out-of-the-window” view for each ship together with the integrated display of electronic map and radar. Based on the input related to interaction design solutions from study 1, a new design layout, F, was developed. The F-layout is based on E, has less information density, is a bit darker, and has clearer symbols for ships, speed, and direction. This design solution includes an additional large screen next to the displays, showing a larger area of the fjord, to better support planning ahead (Braseth et al., 2023).

3.4 Procedure

Both studies took place in a simulator setting. The procedure of both studies was the same. First, participants were welcomed and informed about the project. Then they were informed about the study, their role, and which data we were going to collect. As the scenarios were videotaped in advance, it was not possible for the participants to take actions. Instead, they were asked to act as expert commentators and verbalize their observations and the actions they would have taken if this had been a real situation. They were also asked to verbalize if and when they would like to take manual control of the ships. The different interaction design layouts were presented to the participants for them to familiarize with the set-ups. After signing an informed consent and filling in a background questionnaire, the scenario started. After each scenario, the participants answered questionnaires. Semi-structured interviews were performed.

3.5 Data collection

For both studies, audio and video data for each participant was collected. For evaluating workload, the NASA Task Load Index (NASA-TLX) was used (Hart and Staveland, 1988). This is a widely used subjective multidimensional assessment tool for self-evaluation of workload. The items in NASA-TLX consist of Mental demand; Physical demand, Temporal demand; Performance; Effort; and Frustration.

Three statements related to situation understanding was developed for this study, one for each level of situation awareness (SA) as proposed by Endsley (2013). The statement used for SA1 (necessary data to understand the situation), was: *In this test scenario, the workstation was set up so that the overall traffic picture was always clear to me.* The following statement was used for SA2 (assembling the data into a meaningful picture): *In this test scenario, the workstation was set up so that it was easy to perceive the situation from the perspective of the autonomous vessel(s).* For SA3 (support a projection into the near future), the statement was: *In this test scenario, the workstation was set up so that it was easy to assess the future situation and whether I should intervene and take control of the autonomous vessel(s).*

4 Findings

4.1 Supervising one vs three ships

A significant difference in workload ratings by the participants in study 1 was found for supervising one ship vs. three ships (one-tailed paired t-test, p=0.01), see Figure 2.

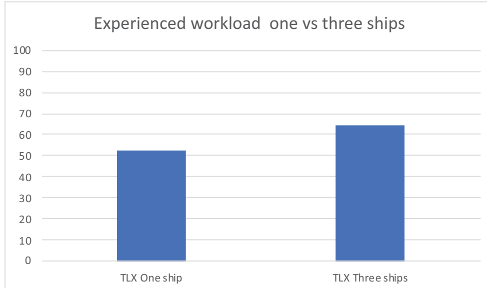

Figure 2: Experienced workload for one ship and three ships (average workload score in study 1).

This finding was supported in the interviews, in that the participants were quite reluctant when asked how many ships they felt it was comfortable to have control over at any given time. Most participants thought it would be possible to have responsibility for three ships simultaneously, but not more than three, some would prefer not more than two. The reason given for this, was that situations that they need to deal with may arise at several ships at the same time, making it quite demanding for the operators. Some stated that when everything is normal, and none of the ships are close to shore, it is possible to supervise more ships. Very few participants answered that they would feel comfortable supervising ships in two different maritime areas, as this would require understanding of the traffic situation, currents in the sea and weather conditions in two different areas simultaneously and thus be very mentally demanding.

4.2 Novices vs. Experts

Novices and experts experienced a similar level of workload in study 1, see Figure 3.

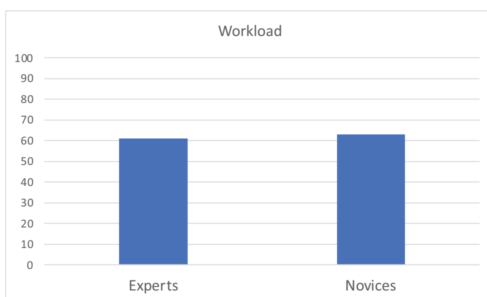

Figure 3: Workload ratings by experts and novices.

When looking more closely at the workload ratings for the individual interaction design solutions, we can see that experts reported a bit higher workload in condition A, while novices reported a bit higher workload in condition D.

Figure 4: Workload ratings by experts and novices in five different design layouts.

In the interviews, the novices expressed that they found layout D particularly challenging, due to a lot of information that needed to be processed simultaneously.

4.3 Design layout for three ships

To investigate how different interaction design solutions impacted workload when supervising three autonomous ships (C, D, E, F), we looked at the workload ratings by the experts only and used data from both study 1 and study 2. Figure 5 present these results.

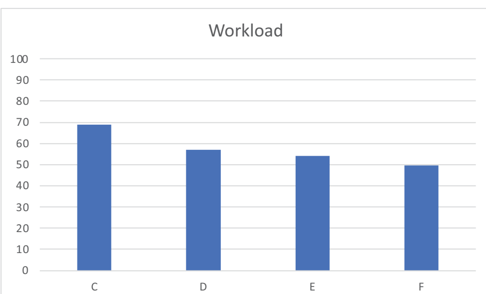

Figure 5: Navigators´ workload while supervising three autonomous ships in four different layouts.

As can be seen in Figure 5, the workload ratings were quite similar across design layouts when supervising three autonomous ships. Layout C is rated a bit higher in workload than the other interaction design solutions.

In the interviews, the navigators stressed that layout C was not to be recommended, as two out of three ships are out of sight for 40 seconds at a time. They expressed that it felt very uncomfortable and unsafe not having continuous overview of the ships they were responsible for.

The three statements related to situation understanding were combined in an average score, see Figure 6. Here it seems like situation understanding was a bit higher for layout F than for the other design layouts while supervising three ships.

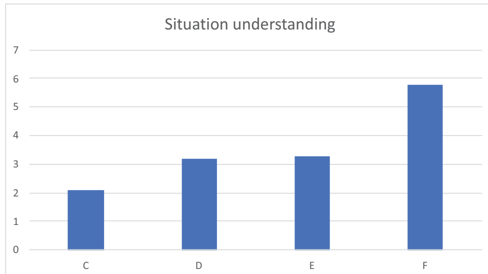

Figure 6: Navigators' reported situation understanding while supervising three autonomous ships in four different interaction design layouts.

In the interviews in study 2, the participants stated that they felt they had better overview when supervising ships with layout F.

In study 1, all participants preferred the setup as in E for supervising three ships. In study 2, all participants preferred the layout as presented in the F condition. What the participants appreciated about the layout in F was that it was easy to distinguish the vessels they were responsible for from other vessels. Furthermore, this display presented necessary and sufficient information for having a good situation understanding at any given time. It was easier to see speed and direction of the ships and they were able to observe different situations much earlier which made it possible to be more proactive with this design layout.

5 Discussion

When developing ROCs for remote monitoring of autonomous ships, it is important to design for all conceivable situations so that the ROC operators´ workload is not too high, but also not too low. Our findings indicate that the workload increases with the increasing number of ships to be supervised.

The participants in our studies did not feel comfortable monitoring more than three ships simultaneously, as they believed it will be challenging to supervise ships in different phases (docking, leaving dock, sailing) safely. If a situation arises with one ship and another ship is to dock or depart, the operator will probably not have the mental capacity to handle this.

Analyzes to uncover different types of simultaneous demanding situations should be carried out in advance of implementing this concept, so that the work processes that are developed support workload in all different situations that the operators have to deal with. If the tasks are distributed reasonable between the various operators in the ROC, such as Wrobel et al. (2021) suggest, the operators will also be able to potentially handle different situations that can arise better and with a manageable level of workload.

It will be relevant to carry out further studies to investigate how work processes can be designed to support workload and safe supervision of autonomous ships. It can be assumed that when a ship is to dock or leave a berth, or if a situation arises during the voyage, an operator will only have responsibility for one ship, while an operator may be responsible for several ships that are in transit under normal conditions. If work processes are designed to align safety management and core tasks, as proposed by Størkesen (2021), it is likely that operator workload does not become too high in the different operational modes.

Experience from aviation indicates that dynamic solutions can contribute to optimal utilization of the operators´ resources, while the workload is continuously adapted to the capacity of the individual operators. Future studies should investigate whether a similar way of working as in aviation may also be relevant for future ROCs.

In the D condition, novices reported higher workload than the experts. The design layout in D presents information from the three autonomous ships in parallel. The electronic map and the radar information is displayed separately for each ship in this setup. Here, it was necessary for the participants to mentally integrate information from these sources, for all three autonomous ships simultaneously. Not surprisingly, novices found this layout particularly challenging. As theories of expertise points out, novices are often overwhelmed by the complexity of the information, while experts have deep knowledge in the field and will more easily understand the information even if it is complex (Ericsson et al., 2006). As the navigators are used to mentally integrating information from the electronic map and the radar displays through years of experience, it is reasonable that it will be less demanding for them to interpret the information for all three ships than for the novices.

Based on the data collected in this study, it may thus appear that having maritime, or navigator experience is an advantage to better perceive and understand the information in the displays, especially in complex conditions.

A high degree of autonomy can lead to fatigue if there are few situations to deal with for the operators in a ROC, while navigators' many years of practical maritime experience may deteriorate the less interventions one needs to do. The study by Veitch and Alsos (2022) indicates that operators who monitor autonomous ships have the same role as those who control conventional ships in ensuring safety. It will therefore be important to ensure relevant competence and perform frequent simulator training for future operators in a ROC so that their competence is maintained.

In our study, it appeared that the navigators experienced the various interaction design solutions with approximately the same level of workload. It is possible that we would have found greater differences in workload for the various interaction design solutions with longer scenarios including several challenging situations to be handled. It is also possible that the operators' rating of workload would have been different if they had been able to intervene where they felt this was necessary. Future studies may therefore be designed with longer scenarios and with possibilities for interventions to better investigate how workload may vary in different situations. However, an interesting finding was that the navigators felt they had a better overview of the situation in layout F, the display that had been developed based on their input through the iterative design process.

Different design solutions and different numbers of ships to be supervised seem to some extend affect workload and situation understanding. In addition to interaction design, it is likely that the way work is organized, and the competence required in the ROCs will also affect the operators' performance, workload and situation understanding.

6 Conclusion and further work

Remote operation is a novel area which is in the need for more research, both related to interaction design, competence requirements, and work process for multi-vessel supervision. With the work carried out in LOAS, we have explored perceived workload and situation understanding when supervising one and three autonomous ships in a limited set of display layouts. The participants rated their workload higher when supervising three ships compared to one ship. Workload did also seem to vary related to some of the interaction design solutions. The participants in the two studies preferred, and seemed to have better overview of the situation, with one of the interaction design solutions (F), and we suggest proceeding further with this concept. Additional areas to be researched, is how alarms should be presented, how interventions should be carried out if needed, and how communication with other ships and with the vessel traffic central should be performed. Such aspects could be investigated to ensure that operators are supported in all tasks they are expected to perform from a ROC with an acceptable level of workload.

The studies presented in this paper have provided useful knowledge that can be used in further work. We recommend continuing an iterative approach to gradually get closer to an interaction design solution that supports safe and efficient supervision of one and several autonomous ships from a ROC.

Acknowledgment

This project is financed by the Research Council of Norway. We would like to thank the participants in both studies for their participation and valuable input to our research.

References

Bainbridge, L. (1983). Ironies of automation. Automatica, 19, 775-779

Braseth, A.O., Kaarstad, M., Høstmark, J.B. (2023) Remote supervision of autonomous ships: principles for interaction display graphics (In prep.) Submitted to ESREL 2023

Chi, M.T.H., Glaser R, Farr, MJ. (1988). The Nature of Expertise. Hillsdale, NJ: L. Erlbaum Ass.

Eitrheim, M.H.R., Skraaning, G. & Hurlen, L. (2019). Operation of multiple reactors from a single control room. HWR-1251. IFE, Norway

Endsley, M. R. (2013). Situation Awareness, In Lee J.D. & Kirlik A. (Eds.), The Oxford Handbook of Cognitive Engineering, Oxford University Press

Ericsson, K.A., Charness, N., Feltovich, P.J., Hoffman, R.R. (2006). The Cambridge Handbook of Expertise and Expert Performance. Cambridge University Press

Hart, S.G., Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In Hancock, P.A., Meshkati, N. (eds.). Human Mental Workload. Advances in Psychology. Vol. 52. Amsterdam: North Holland. pp. 139–183.

Hetherington, C., Flin, R. & Mearns, K. (2006) Safety in shipping: The human element. J. Safety Research, 37, 401-11

IMO (2018). IMO takes first steps to address autonomous ships. Accessed 20. Apr. 2021, https://www.imo.org/en/MediaCentre/PressBriefi ngs/Pages/ 08-MSC-99-MASS-scoping.aspx

IMO (1996). STCW Convention. International Convention on Standards of Training, Certification and Watchkeeping for Seafarers, 1978, as amended in 1995. International Maritime Organisation, London.

Kaarstad, M., Braseth, A.O. (2020). Operating autonomous ships remotely from land-based operation centers: The current state-of-the-art. IFE/E-2020/008

Lützhöft M, Nyce J (2014) Integration work on the ship's bridge. J Marit Res 5(2):59–74

Lützhöft M (2004) The technology is great when it works: Maritime technology and human integration on the Ship's bridge. (ph.D. thesis). Linköping University

Porathe, T., Hoem, Å., Rødseth, Ø., Fjørtoft, K., Johnsen, S.O. (2018). At least as safe as manned shipping? Autonomous shipping, safety and 'human error,' Safety and Reliability – Safe Societies in a Changing World, Haugen et al. (Eds), Tyler & Francis Group, London, UK

Reling, T., Lützhöft, M., Ostnes, R., Hildre, P. (2018). A human perspective on maritime autonomy. Augmented Cognition, Springer Int. Publishing.

Størkersen, K. V. (2021). Safety management in remotely controlled vessel operations. Marine Policy, 130, 104349.

Veitch, E., & Alsos, O. A. (2022). A systematic review of human-AI interaction in autonomous ship systems. Safety science, 152, 105778.

Wróbel, K., Gil, M., & Chae, C. J. (2021). On the influence of human factors on safety of remotely-controlled merchant vessels. Applied Sciences, 11(3), 1145.

Zelinski, S., Lai, C. F. (2011). Comparing Methods for Dynamic Airspace Configuration. IEEE/AIAA Dig. Avionics Systems Conference.