TransNav
http://www.transnav.eu
the International Journal on Marine Navigation and Safety of Sea Transportation
Volume 17
Number 4
December 2023
DOI: 10.12716/1001.17.04.23

How Does Maritime Situation Awareness Depend on Navigation Automation and Mental Workload? A Sea Simulator Experiment

G. Müller-Plath, J. Lehleitner, J. Maier, J. Silva-Löbling, H. Zhang, X. Zhang, & S. Zhou
Technische Universität Berlin, Berlin, Germany

ABSTRACT: A good situation awareness (SA) of the navigator is essential for the safety of the ship, especially in coastal areas. In this study, the Unity 3D engine was used to simulate the navigation of a coastal trading vessel along predefined routes in the Baltic Sea. The SA of the helmsman, who was either Chinese or European, was assessed several times with the SAGAT test (Endsley, 1995b, 20-21) and compared between low and high workload conditions and between manual and autopilot navigation. High workload and automated navigation both reduced SA significantly and in an additive manner. No difference was found between Chinese and European participants. In contrast to previous accident analyses of SA, we found that SA level 3 (projection of future states) was most strongly affected by both factors, while SA levels 1 (perception of relevant information) and 2 (comprehension of the current situation) suffered to a lesser extent. Further research is needed to establish specific relationships between types of automation on ships, types of workload, and SA problems in order to design countermeasures.

1 INTRODUCTION

The vast majority of maritime accident analyses identify human error as a main or contributing cause. Recent estimates range from 60 % (Butt et al., 2013) to 89.5 % (EMSA 2022, p. 5). However, according to the classical accident pyramid, accidents are only the tip of an iceberg, with incidents at the visible base and a large amount of near misses and unsafe acts under the surface (William 1959; Grech, Horberry & Koester, 2008, p. 17).

Human factors research could attribute unsafe acts in many cases to a lack of situation awareness (SA), a concept derived from aviation and elaborated by Endsley (1995a). Endsley defines SA as a mental process consisting of three successive levels: the perception of relevant elements in the environment within a volume of time and space (level 1), the comprehension of their meaning (level 2), and the projection of their near-future state (level 3). In the maritime domain, 71% of human error in accident reports (Grech, Horberry & Smith, 2002), or correspondingly 50% of all accidents (Stratmann & Boll 2016) were found to be related to a SA problems. Loss of level 1 SA (perception) was reported most frequently.

The human factor has long been neglected in the maritime domain and has received increasing attention only in recent years (Grech et al. 2008). A focus of the latest empirical research is human decision-making, mostly in the context of developing maritime collision avoidance support systems (e.g. Aylward et al. 2022, Butler et al. 2022, Fan et al. 2023, Kizilay et al. 2023, Kartoglu et al. 2022), applying or citing a variety of methods, including ship simulators, interviews, surveys, observations, case studies, and even SA questionnaires (e.g. Clemente et al. 2014). In this type of research, SA is one of several psychological factors that determine decision making. However, what in turn are the determinants of SA?

Endsley & Jones (2011, ch. 3) suggested eight, mostly interacting, factors that may impair SA at different levels: Attentional tunnelling, memory failures, stress due to anxiety, fatigue etc., mental overload, misplaced salience/distraction, creeping complexity, errant mental models, out-of-the-loop syndrome due to automation (Wickens 2002, Parasuraman et al. 2008). Although most of them have been identified in analyses of maritime accident reports (see e.g. Stratmann & Boll 2016), to our knowledge, it has not been empirically demonstrated in the maritime domain that one or more of these factors causally affects SA at a subthreshold level, i.e. without a subsequent incident or accident (see the current literature review by Fan at al., 2023). The present pilot study is a first step towards filling this gap

In a bridge simulator experiment, we tested the SA of the helmsman of a coaster vessel during navigation while manipulating two of the aforementioned factors: high versus low mental workload, and automated versus manual navigation. The graph in Figure 1 visualises the proposed hypotheses in the 2 x 2 experimental design: (H1) Main effect of workload: Averaged over the two levels of automation, high workload was expected to reduce SA compared to low workload, due to cognitive overload (Wickens 2002; Parasuraman et al. 2008). (H2) Main effect of automation: Averaged over the two levels of workload, automated routing and steering was expected to reduce SA compared to manual, due to the out-of-the-loop syndrome (the operator loses vigilance because his mental presence is not needed in operating the ship, Endsley & Kiris 1995). (H3) No interaction: As automation reduces workload (Parasuraman et al. 2008; Endsley & Kaber 2011), there could be an interaction between automation and workload such that the impact of high workload on SA is less severe in the automated than in the manual condition. Due to a lack of literature on this, we do not claim such an interaction a priori.

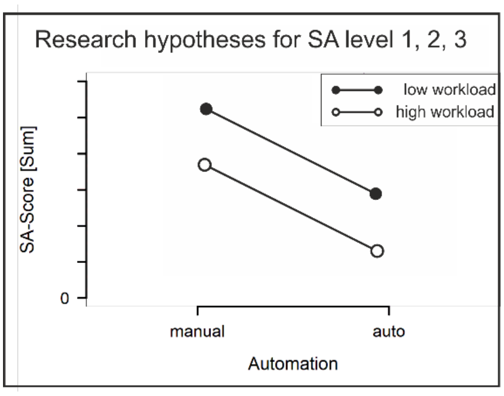

Figure 1. Hypothesised SA of the helmsman in navigating a simulated coaster under conditions of low and high mental workload, and manual versus automated route planning and steering.

2 METHOD

2.1 *Participants*

16 students of the Technische Universität Berlin (23-33 years of age, 8 female/8 male) participated as volunteers. In order to ensure intercultural validity, 8 were European (fluent German-speaking), and 8 were Chinese (native language Mandarin). All participants had normal or corrected-to-normal vision and hearing, no neurological impairments, no medication that impairs driving ability, and did not possess a recreational boat license or nautical patent. Since they were not familiar with maritime navigation, they conducted a training session before taking part in the two test sessions, which are described in section 2.3 below. They were paid 40 Euros for participation.

2.2 *Simulator*

The sea simulator was programmed in our lab with the Unity 3D engine (Release Unity2019.1.6f1). The participant was shown the Point of View (POV) from the bridge of a coaster trading vessel with overall length 75 m, width 12 m, draught 4 m, projected onto a screen of 2 x 3 m (see Fig. 2). At the top of the screen, some navigation relevant data were displayed (water depth, speed, current course in degrees, and elapsed time). The participant was seated in a mockup ship bridge at a distance of 3.5 m to the screen and operated the ship's movement with a joystick, and the POV with a mouse. The area to be navigated was a coastal region of the Baltic Sea, the Isefjord in Denmark (see Fig. 3), in good weather condition. It contained several harbours and marinas, isles and islets, a buoyed fairway, and some narrow anchorages. Water depth was on average 5 - 7 m, but shallow near the coast, in harbour entrances, and at some single spots marked with cardinal buoys. The water depths in the simulator corresponded exactly to those indicated in the chart with linear interpolation in between. Other navigation aids included buoys, harbour buildings with sailboat masts, and coastlines with trees and cylindrical towers. Also, a variety of other ships were visible consisting of four distinct types: coaster, ferry, sailing boat, motor boat. A tablet was used to simulate the ECDIS with the marked route, distance to the next waypoint, and Automatic Identification System (AIS) icons for other ships in proximity (red: within a critical distance of 0.5 nautical mile (nm), green: outside a radius of 0.5 nm; see Fig. 7).

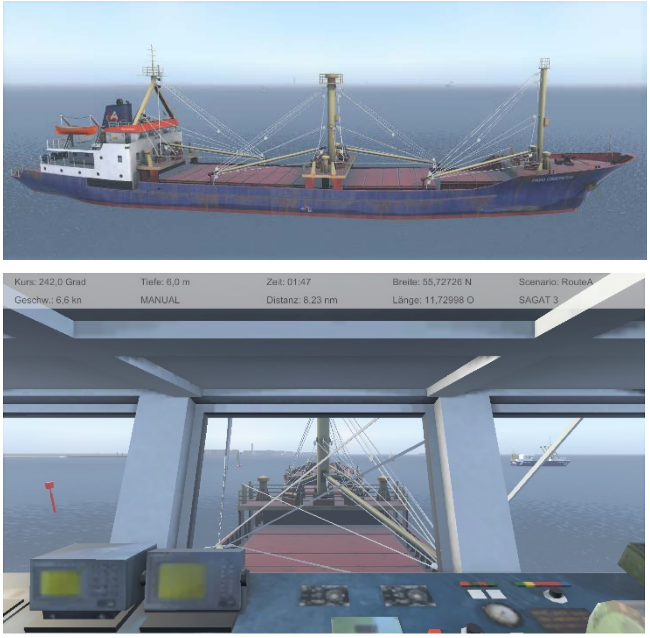

Figure 2. Unity model of the coaster to be navigated through the Isefjord in Denmark, as seen from external (above) and from the bridge (below, point-of-view of the participants).

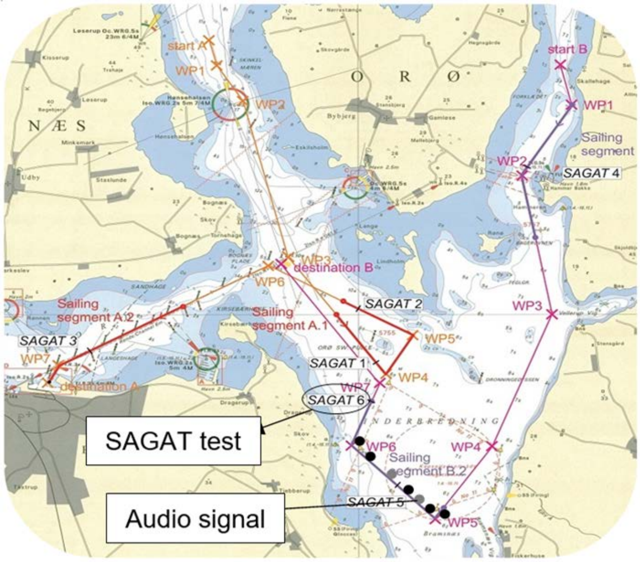

Figure 3. Simulated test area: Isefjord in the Baltic Sea, Denmark, with the two routes (A and B), waypoints (WP), the segments to be sailed, and the six SAGAT test points (three on each route). The position of SAGAT test 6 and the positions of preceding sound signals are marked as examples. participants).

2.3 *Procedure*

Participants were tested singly in the simulator for about two hours. Smooth communication was guaranteed by matching the spoken language of experimenter and participant (Mandarin or German).

As our participants were not familiar with maritime navigation, they started with a training session consisting of the following tasks: Steer a course of 000 degrees (north), round a red buoy, head for a green buoy, speed up and slow down in forward and reverse, stop the ship, explain displays. They were also presented with the audio signals relevant for their workload condition (see Table 2), to which they had to respond with the appropriate action. Finally, a full SAGAT test (Table 1) was conducted.

In two subsequent test sessions, participants navigated two routes that were set up in advance on the "ECDIS" (the tablet). Each route consisted of seven waypoints and had a total length of 7.5 nm of which two segments of 1.5 and 2 nm each had to be actually sailed (see Figure 3). There was a 15 minute break between the two routes..

Situation Awareness (SA) was tested three times along each route at unpredictable points. As a measurement tool, we applied the Situation Awareness Global Assessment Technique (SAGAT) developed by Endsley (1995b, 2021). It uses the "freeze technique", where a situation is stopped and the participant answers questions at all three levels of SA. Here, the simulator was set to freeze and masked, seven questions were asked (see Table 1) and the answers were recorded along with the true situation. Two questions referred to SA level 1 (perception), two to level 2 (comprehension), and two to level 3 (projection). The questions are shown in Table 1.

An additional seventh question, which was always asked first, was a control question that had no direct relation to ship navigation, but served to control for the specificity of our navigation-related SAGAT: we expected the six SAGAT questions to be affected by automation and workload, as hypothesised in Fig. 1, but not the control question. Some questions could only be solved by keeping a good lookout, while others required constant attention to the "ECDIS" or "bridge instruments" at the top of the screen.

The automation of navigation was manipulated by the use of an autopilot: On one route, the participant steered the ship manually (the waypoints marked on the "ECDIS" had to be reached, and an SOG of 8 and 10 knots respectively was recommended for the two segments), while on the other route the ship was navigated by the autopilot at those SOGs. Every participant started with the same route, but the mapping of the two routes to the two levels of automation was randomised across participants.

Mental workload was manipulated using audio signals indicating specific secondary tasks, which appeared at irregular intervals before and after the SAGAT tests. Participants had to complete the task immediately after the audio signal. Low workload was induced by presenting only one of two possible audio signals, linked to simple tasks. To induce high workload, two more signals with more complex tasks were added and applied 2-4 times between SAGAT tests (see Table 2). While automation of navigation was manipulated within participants (each participant performed both levels), mental workload was implemented as a between-subjects factor, with n = 8 participants in the low and n = 8 in the high workload group (4 Chinese, 4 German speaking each).

Table 1. SAGAT questions 1. – 6., source of information, and SA level, as asked three times along each of the two routes. The first question 0 served as a control question unspecific to maritime SA.

Question	Source	SA Level
0. What man-made objects or structures are visible on the coast ahead?	lookout	(control)
1. How many vessels are currently in close range? (AIS symbols in red)?	"ECDIS"	1
2. What type of vessel are they (sailing boat, small motorboat, coaster, ferry)?	lookout	1
3. What is your current position? (Draw in the other tablet)?	"ECDIS"	2
4. How much water do you currently have under your keel (in m/cm)?	depth display	2
5. What is the distance to the next waypoint?	"ECDIS"	3
6. How will you change course there? (Indicate in degrees and whether to port/left or starboard/right.)	"ECDIS"	3

Table 2. Audio signals and tasks to induce low/high workload. Low workload: 1 signal from S1, S2 between every two SAGAT points. High workload: 2-4 signals from S1, S2, S3, S4 between every two SAGAT points.

Audio signal	Task to be executed
S1 short beep, middle frequency	do nothing
S2 sequence of beeps ("fuel alarm")	cancel with button press
S3 sequence of 3 short beeps	rotate your POV by 360° using the mouse
S4 cell phone ring tone	memorize a given 4-digit number and recall it after a minute (new beep)

2.4 *Data Analysis*

At each SAGAT query (three per route), the real data of the situation with regard to the seven questions were recorded at the moment the participants left the situation (the simulator was "frozen"). Dependent on the impact of errors for safe navigation, the accuracy of each answer was scored with 0, 1, or 2 points and summed up across the three tests of each route. The scores were averaged within each level and across all levels. Thus, for each of the two routes, i.e. for the automated and the manual navigation condition, each participant received six scores for the six questions, three average scores for SA level 1, 2, 3, and a total SA score, all ranging from 0 to 6 points. The statistical analysis of the data followed the procedure described in the next section.

3 RESULTS

Figure 4 shows the mean total SA scores of the 8 participants of the high workload group and the 8 participants of the low workload group, with manual and automated navigation each. Figure 5 shows the same separately for the three levels of SA and the six individual SAGAT questions. In order to generalise from the small sample of participants to a population, we conducted statistical significance tests of the two main effects (hypotheses H1, H2) and the interaction (H3) of the factors workload and automation in the 2 x 2 variance analytic design for each of the ten graphs shown in Figures 4 and 5. Since hypotheses H1 and H2 were formulated as directed effects and the factor automation was a within-subjects factor, these hypotheses were tested with one-tailed t-tests for two independent samples instead of variance analytic F-tests in order to maximise statistical power. (The independent samples t-test on the appropriate sum or difference scores of the within-factor, respectively, is equivalent to testing the main effects or the interaction in a mixed 2 x 2 ANOVA; the homogeneity of variances is not critical here because the two groups have equal size; Posten 1984.)

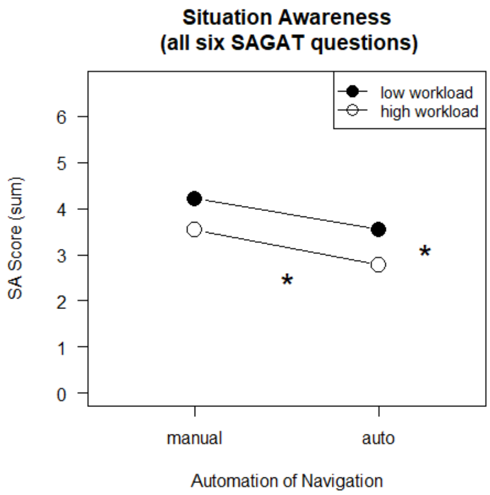

Figure 4. Mean total SAGAT scores (range 0-6) under manual and automated navigation for the two groups of participants with low workload (black dots, n = 8) and high workload (white dots, n = 8). The stars indicate statistically significant ($p < 0.05$) main effects.

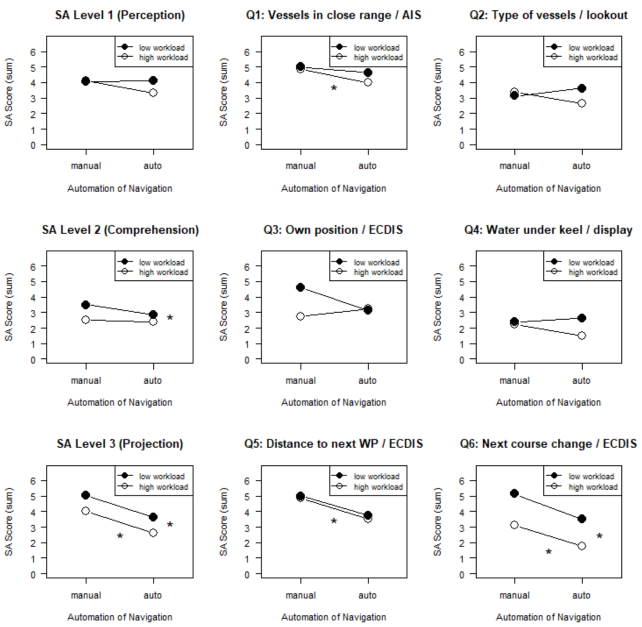

Figure 5. Left column of graphs: Mean SAGAT Level 1, Level 2, and Level 3 scores (range 0-6) under manual and automated navigation for the two groups of participants with low (black dots, n = 8) and high workload (white dots, n = 8). Middle and right columns of graphs: Mean SAGAT scores (range 0 – 6) for the individual questions, see Table 1. The stars indicate statistically significant ($p < 0.05$) main effects.

The results graph in Fig. 4 looks almost exactly like the hypotheses graph (Fig. 1). The two hypotheses H1 and H2 were also confirmed statistically: Averaged over the two conditions of automation, high workload reduced SA compared to low workload (t = 2.269, df = 14, p = 0.019). Likewise, averaged over the two workload conditions, autopilot navigation reduced SA compared to manual navigation (t = 3.914, df = 14, p = 0.0008). However, despite its significance, the effect of automation does not seem very large. Further insight is gained from a detailed look at the SA levels and the individual questions in Fig. 5: The proposed pattern of results was most evident only in SA level 3, projection of perception into future states. Here, automation and workload had a strong and significant impact (t = 4.817, df = 14, p = 0.0001 for the directional main effect of automation; t = 1.784, df = 14, p = 0.048 for the directional main effect of workload), and both questions were affected. SA level 2, comprehension of the current situation, was significantly impaired only by the high workload (t = 1.829, df = 14, p = 0.044 for the directional main effect of workload) but not by the automated navigation (t = 0.732, df = 14, p = 0.238 for the directional main effect of workload). This seems to be due to the fact that the latter effect was not consistent in the two questions of this SA level. The most basic SA level 1, perception of relevant information, was not significantly affected by either of the factors (t = 1.417, df = 14, p = 0.089 for the directional main effect of automation; t = 0.73, df = 14, p = 0.249 for the directional main effect of workload). However, it might be worth noting that the performance on one of the test items, the question of how many other vessels were in close range, was significantly impaired by the autopilot navigation (t = 2.054, df = 14, p = 0.029 for the directional main effect of automation).

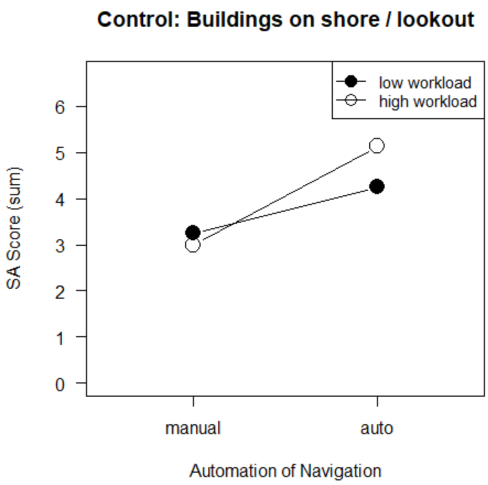

Figure 6. Mean scores (range 0 – 6) in the control question. Symbols as in Fig. 4.

A control question was included in the SAGAT (Table 1, first row) to check whether workload and navigation automation affect on-board perception in general or navigation-related situational awareness in particular. Interestingly, as Figure 6 shows, the result of this test item was completely different from all the "real" SAGAT items: Buildings on shore were discriminated even better under autopilot navigation than under manual navigation. There was a small positive effect of high workload. (As we had no hypotheses for this item, no significance tests were carried out.)

Further analyses showed that neither the native culture of the participants, Chinese or European, nor the mapping of automated and manual navigation on the two routes A and B, had a statistically significant effect on the SA results.

4 CONCLUSION

Human error, mostly due to a lack of situation awareness (SA), is responsible for the vast majority of all accidents at sea. Maritime human factors research sees human error not as the end of an investigation, but as the starting point. It is not the cause of a problem but the effect of a deeper trouble, a sensor indicating that something is wrong in the human-machine system. Reason (2000) coined the famous Swiss cheese model of human error which states that a variety of latent and active failures by various actors, from individual human operators to designers and economic and legal organisations, must align like holes in a Swiss cheese in order to produce an accident. The small proportion of actual accidents relative to latent unsafe human acts (accident pyramid) is illustrated herein by the low probability of multiple holes in a cheese being aligned.

To date, most research on the determinants of maritime SA has focused on accident reports. This has mostly identified problems at SA level 1 (perception). Our research has taken the first step in a different direction: We investigated SA problems at a subthreshold level, where no accident occurred. In order to causally examine the detrimental effects of high workload and of an out-of-the-loop state due to automated navigation, we conducted a controlled experiment in a sea simulator, simulating coastal navigation. We confirmed harmful effects of both factors, with no interaction between them. That is, the effects of workload and automation were additive, with the worst SA under high workload and automated navigation. The idea that automation might compensate for workload by reducing it was not supported. Furthermore, in contrast to the analysis of accident reports, we found the most severe reduction in SA not at SA level 1 (perception) but at level 3 (projection). It is not surprising that this result does not emerge from accident analyses: When people are simply asked what they did or did not perceive, only SA level 1 is addressed. Almost no one is able to report that he or she was unable to project their perception into the future. The SAGAT test applied here allows a much deeper insight into the cognitive processes underlying SA loss. Despite some drawbacks mentioned below, Burmeister et al. (2021, p. 22-23) pointed out that the SAGAT is the most valid test of SA available.

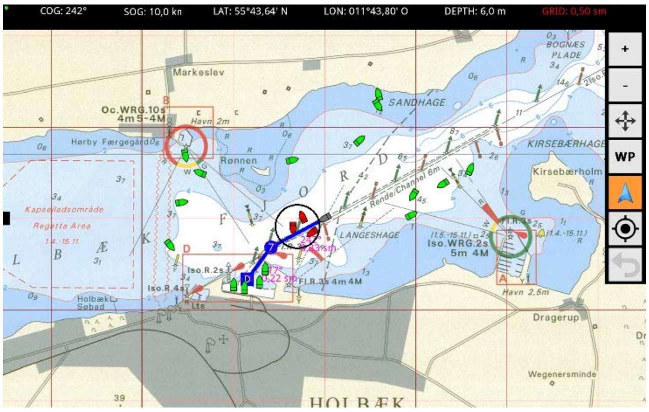

Figure 7. Screenshot of the simulated "ECDIS / AIS" display at one SAGAT test point ("freeze"). In the simulated AIS, vessels in close range, defined as a radius of 0.5 nm, were displayed as bright red icons, other vessels as green icons. The black circle, which was not part of the original display, shows the red icons for reasons of clarity in the black-and-white print.

Although SA level 1 (perception) was not much affected by our manipulation of workload and automation, one result is worth highlighting: In our SAGAT questions 1 and 2, participants had to indicate how many other vessels were in close range and of what type. This is an everyday problem in coastal navigation, as vessels of many types (ferries, motor yachts, sailing boats and other fishing or trading vessels) frequently cross fairways and harbour entrances. In our simulation, the AIS icons of vessels in close range were displayed in bright red (Fig. 7). We found a significant decrease in Question 1 scores when sailing on autopilot compared to manual. We argue that this effect was indeed a consequence of the out-of-the-loop syndrome, as our control question (indicate how many man-made structures are visible on shore) was answered even better when sailing on autopilot.

There are two limitations to our study, apart from the fact that it is only a simulation. First, the SAGAT technique for measuring SA suffers from the problem that it has to be repeated several times to give reliable results, but then the participant knows the questions to be asked and can prepare mentally. However, this implies that participants have an unrealistically good SA because of this preparation, and in reality SA problems may be even worse. This problem can only be addressed with a much larger number of participants, each of whom takes only one SAGAT. Secondly, our participants were students, not professional officers at sea. However, in our previous research on maritime navigation in the simulator and at sea, we found no systematic differences between the two populations (Müller-Plath et al. 2018; Müller-Plath 2019), which is probably due to fundamental laws of human-machine interaction being involved.

Future research should firstly validate the results with experts in maritime navigation and a larger sample. Second, it should extend the line of study outlined here and investigate more specific questions such as what kind of automation affects what kind of perceptions, understandings, and projections in which way? And thirdly, how can training or tools be designed to counteract the loss of SA, as simply more automation does not seem to be the "silver bullet"?

ACKNOWLEDGEMENTS

The authors would like to thank Richard Gross for programming the sea simulator and solving many technical problems in a short time.

REFERENCES

Aylward, K., Weber, R., Lundh, M., MacKinnon, S.N., & Dahlman, J. 2022). Navigators' views of a collision avoidance decision support system for maritime navigation. The Journal of Navigation, 1-14.

Butler, G.L., Read, G.J., & Salmon, P.M. 2022. Understanding the systemic influences on maritime pilot decision-making. Applied Ergonomics, 104, 103827.

Burmeister, H.-C., Grundmann, R., Hochgeschurz, S., Hohnrath, P., Ujkani, A. 2021. Increasing Maritime Situational Awareness by Augmented Reality Solutions. Hamburg, Germany: Fraunhofer Center for Maritime Logistics and Services CML

Butt, N., Johnson, D., Pike, K., Pryce-Roberts, N., & Vigar, N. 2013. 15 Years of Shipping Accidents: A review for WWF. Southampton Solent University. http://awsassets.panda.org/downloads/15_years_of_shipping_accidents_a_review_for_wwf_.pdf. Last access 08 April 2023.

Clemente, S., Loia, V., & Veniero, M. 2014. Applying cognitive situation awareness to collision avoidance for harbour last-mile area safety. Journal of Ambient Intelligence and Humanized Computing, 5, 741-745.

European Maritime Safety Agency (EMSA) 2022. Annual Overview of Marine Casualties and Incidents 2021. https://www.emsa.europa.eu/we-do/safety/accident-investigation/download/6955/4266/23.html. Last access 08 April 2023.

Endsley, M.R.. 1995a. Toward a theory of situation awareness in dynamic systems. Human Factors 37(1): 32-64.

Endsley, M.R. 1995b. Measurement of situation awareness in dynamic systems. Human Factors 37(1): 65-84.

Endsley, M.R. 2021. A systematic review and meta-analysis of direct objective measures of situation awareness: A comparison of SAGAT and SPAM. Human Factors, 63(1), 124-150.

Endsley, M.R. & Jones, D.G. 2011: SA Demons: The Enemies of Situation Awareness. In: Designing for Situation Awareness: An Approach to User-Centered Design, ch. 3, pp. 31–41. Boca Raton: CRC Press.

Endsley, M.R. & Kaber, D.B. 1999. Level of automation effects on performance, situational awareness and workload in a dynamic control task', Ergonomics, vol. 42, no. 3, pp. 462-492.

Endsley, M.R. & Kiris, E.O. 1995. The Out-of-the-Loop Performance Problem and Level of Control in Automation. Human Factors, vol. 37, no. 2, pp. 381–394.

Fan, S., Blanco-Davis, E., Fairclough, S., Zhang, J., Yan, X., Wang, J., & Yang, Z. 2023. Incorporation of seafarer psychological factors into maritime safety assessment. Ocean & Coastal Management, 237, 106515.

Grech, M. & Horberry, T. 2002. Human error in maritime operations: Situation awareness and accident reports. Paper presented at the 5th International Workshop on Human Error, Safety and Systems Development, Newcastle, Australia.

Grech, M. R., Horberry, T. J., & Koester, T. 2008. Human factors in the maritime domain. Boca Raton, FL: CRC Press

Kartoglu, C., Senol, Y.E., & Kum, S. 2022. Situational awareness of navigators in high-speed craft bridge navigation operations. Australian Journal of Maritime & Ocean Affairs, 1-15.

Kizilay, F.E., Arslan, O., Akyuz, E., & Kececi, T. 2023. Prediction of human error probabilitiy for officers

during watchkeeping process under SLIM approach. Australian Journal of Maritime & Ocean Affairs, 1-18.
Müller-Plath, G., Jung, D. & Müller, M. 2018. Research-based design and usability guidelines for electronic charting systems (ECS) in yachting and boating. International Journal of e-Navigation and Maritime Economy, 10, 32-48.
Müller-Plath, G. 2019. How does digital navigation on sailboats affect spatial abilities at sea?. TransNav: International Journal on Marine Navigation and Safety of Sea Transportation, 13(2), 299-310.
Parasuraman, R., Sheridan, T.B., & Wickens, C.D. 2008. Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. Journal of cognitive engineering and decision making, 2(2), 140-160.
Posten, H. O. (1984). Robustness of the two-sample t-test. In: D. Rasch & M.L. Tiku (eds.). Robustness of statistical methods and nonparametric statistics: 92-99. Dordrecht, Holland: Kluwer Academic Publishers Group.
Reason, J. (2000). Human error: models and management. BMJ, 320, 768-770.
Stratmann, T. C. & Boll, S. 2016. Demon Hunt-The Role of Endsley's Demons of Situation Awareness in Maritime Accidents. In 6th International Conference on Human-Centred Software Engineering (HCSE)/8th International Conference on Human Error, Safety, and System Development (HESSD) (pp. 203-212). Springer International Publishing.
Wickens, C. D. (2002). Situation awareness and workload in aviation. Current directions in psychological science, 11(4), 128-133.
William, H. H. 1959. Industrial accident prevention: A scientific approach. Boston: McGraw-Hill Book Company.