TRANSNAV
http://www.transnav.eu
the International Journal
on Marine Navigation
and Safety of Sea Transportation
Volume 16
Number 2
June 2022

DOI: 10.12716/1001.16.02.13

EEG Based Workload and Stress Assessment During Remote Ship Operations

R. Kari[1], A.H. Gausdal[2] & M. Steinert[3]
[1] *Norwegian University of Science and Technology, Alesund, Norway*
[2] *Kristiania University College, Oslo, Norway*
[3] *Norwegian University of Science and Technology, Trondheim, Norway*

ABSTRACT: Autonomous and remotely controlled ships present new types of human factor challenges. An investigation of the underlying human factors in such operations is therefore necessary to mitigate safety hazards while improving operational efficiency. More tests are needed to identify operators' levels of control, workload and stress. The aim of this study is to assess how increases in mental workload influence the stress levels of Shore Control Centre (SCC) operators during remote ship operations. Nine experiments were performed to investigate the stress levels of SCC operators during human-human and human-machine interactions. Data on the brain signals of human operators were collected directly by electroencephalography (EEG) and subjectively by the NASA task load index (TLX). The results show that the beta and gamma band powers of the EEG recordings were highly correlated with subjective levels of workload and stress during remote ship operations. They also show that there was a significant change in stress levels when workload increased, when ships were operating in harsh weather, and when the number of ships each SCC operator is responsible for was increased. Furthermore, no significant change in stress was identified when SCC operators established very high frequency (VHF) communication or when there was a risk of accident.

1 INTRODUCTION

Autonomous ships that can be operated remotely have been envisaged as both safer and as a way of improving maritime operational efficiency while reducing crew-related costs. Several developmental and research projects on this topic are therefore being conducted globally. This technology is still in its infancy, and more knowledge about its operation is required. In remote ship operations, officers are relocated from onboard the ship to Shore Control Centers (SCCs). Technical autonomous ship controllers (ASCs) are placed onboard the ship to allow SCC operators to connect and interact with onboard control systems [1]. The SCC operational modes are a combination of monitoring and control modes [2]. Generally, SCC operators monitor status indicators for weather, location, collision, visibility, engine and propulsion. The control modes include status investigation, ASC updates, remote operation and intervention [3].

However, introducing new approaches to control ships remotely also introduces different types of human factor challenges from those found in traditional maritime systems, with regard to both human-machine and human-human interactions [4]. As a result, the 103rd session (5-14 May 2021) of the Maritime Safety Committee has approved the outcome of a regulatory scoping exercise for the use of Maritime Autonomous Surface Ships (MASS) [5]. At that session, terms such as master, responsible person, crew, remote control centers and remote operators as seafarers were identified as potential gaps in the operation of MASS, which should be addressed before extensive deployments of autonomous ships take place [5]. An investigation of the human factors underlying remote ship operations is therefore necessary in order to mitigate safety hazards while improving operational efficiency.

If the hypothesis that the human-machine interface (HMI) can be successfully implemented is confirmed, it is expected that SCC operators will control and monitor up to six ships simultaneously [6]. These operators will require appropriate levels of control, situational awareness and workloads. To find out what the appropriate levels are, a quasi-experimental project, MUNIN, has tested the hypothesis with data from SCC and maneuvering systems. The results indicate that the hypothesis that HMI can be successfully implemented should be accepted; however, tests of the remote maneuvering system were not fully successful [6]. More tests are therefore needed, and the aim of the current study is to assess how increases in mental workload influence the stress levels of SCC operators during remote ship operations.

To achieve this aim, we first performed a literature review to investigate the human factors which influence monitoring operations. The results of the review were then used to develop a series of hypotheses to (i) identify which types of variables (ship indicators) affect workload during monitoring operations, (ii) verify that workload and stress affect monitoring operations, and (iii) identify whether brain signals captured by electroencephalography (EEG) can be utilized to assess the stress levels and workloads of SCC operators during remote ship operations. Finally, two SCC experiments were performed to analyze low and high workload scenarios.

The remainder of this paper is organized as follows: Section 2 presents the literature review and hypotheses; Section 3 presents the material and methods; Section 4 presents the results of the experiments; Section 5 discusses the results; and Section 6 concludes this study and presents a roadmap for future research.

2 LITERATURE REVIEW

2.1 *Remote Ship Operations*

The Maritime Safety Committee of the International Maritime Organization (IMO) approved interim guidelines for MASS trials in 2019 that defined four degrees of ship autonomy. The first degree of ship autonomy includes ships with automated processes and decision support. Onboard seafarers operate and control shipboard functions and systems on ships with the first degree of autonomy. The onboard crew are ready to take control of automated and unsupervised operations [7]. The second degree of ship autonomy includes ships which are controlled remotely by onboard seafarers. On ships of this degree, the ship is operated and controlled from a distant location, but there are also crew onboard the ship who can take control of shipboard systems and functions [7]. The third degree of ship autonomy, the ship is remotely controlled without any seafarers on board: as with the second degree, the ship is controlled from another location, but in this case there are no crew on board. The fourth degree of ship autonomy includes fully autonomous ships which can make decisions and determine the actions to be taken by themselves [7].

It is important to mention that the operation of an autonomous ship can involve a combination of one or more control modes and levels of autonomy during a voyage [2, 7, 8]. For example, operators in the SCC can employ direct remote control when a ship approaches port traffic, in harsh weather or in unexpected traffic situations [9]. Hence, in a ship-shore system with any level of automation, operators (humans) are still involved, but are distributed in SCCs instead of operating conventionally onboard ships [7, 9].

2.2 *Human SCC Operators*

Human SCC operators are defined as officers of the watch who are responsible for monitoring the ship and intervening if necessary [2]. According to the MUNIN project, SCCs will be responsible for most supervisory monitoring and control operations [7]. In the course of a voyage, operators' dynamic navigation tasks are comprised of different aspects, such as: (i) planning the mission, confirmation and designation; (ii) handling critical situations during the voyage; (iii) monitoring the ship's status and health, judging whether the ship needs maintenance and preparing a maintenance plan if necessary [10]; (iv) communicating with other ships and shore elements; (v) maneuvering the ship in ports and waterways, either remotely or from on board; and (vi) gaining experience and learning from the outcomes of operations to improve future activities. Accordingly, the operator's performance depends on three factors: problem recognition, making correct and timely decisions, and acting correctly continuously and on demand [10].

Investigations have clearly indicated that human errors cause the majority of maritime accidents, and this highlights the importance of human factor studies. However, the main question is how human factors should be studied, since human errors do not occur in an isolated environment. Indeed, human errors are intermixed with other problems such as the complexity of human interactions, including human-human interactions and human interactions with other factors in the system [11]. Therefore, several studies [4, 12, 13] have investigated human factor issues that could affect human-human and human-machine interactions during remote ship operations and within SCCs. These studies have revealed mental workload and stress as the human factors with the highest impact on human errors.

2.3 *Mental Workload and Stress*

The mental workload caused by the various challenges of modern shipping, including complex systems, high levels of automation and decreasing crew sizes, has been identified as the main human factor affecting human performance in this context.

This mental workload is cognitive or perceptual and is caused by the amount of mental effort which an operator must expend to perform a task or a series of tasks [11]. Kari et al. [4] have identified that SCC operators with workloads that are too demanding may have difficulties understanding the situation of the ship they are monitoring. Generally, the best operator performance occurs at an intermediate level of mental workload [11].

The operator's stress level is related to situations in which the operator perceives that the available resources are insufficient to manage the task and situation. High levels of stress can lead operators to focus on limited aspects of their tasks and overlook other aspects. As a result, high levels of stress can lead operators to take unsafe and risky actions [11]. This means that a perceived mismatch between the demands of a task or event and an individual's resources leads to an increase in stress levels [14]. Moreover, several studies have indicated that stress and mental workload are strongly interconnected. For instance, it has been found that there is a positive correlation between mental workload and stress, which implies that when operators are exposed to greater workloads their stress levels tend to increase [14].

2.4 *Related Work*

Dussault et al. [15] have studied the effect of mental workload without exposing participants to actual physical risk by using EEG and ECG to investigate the cortical and cardiovascular changes which occur during simulated flight. A total of 12 pilots participated in the experiment, which involved 10 sequences with different mental workloads. The results indicated that theta band power was lower at the central, parietal, and occipital regions of the brain during the two simulated flight rest sequences than it was during visual and instrument flight sequences. In addition, rest sequences resulted in higher beta (at the C4 region) and gamma (at the central, parietal, and occipital regions) band powers than active segments did. In another study, Qing et al. [16] investigated mental workload during the production process by using EEG and Galvanic Skin Response (GSR). Participants were divided into two groups according to whether they were novices or veterans. The novice participants had higher levels of mean voltages in the right hemisphere of their brains for SMR, theta, beta and gamma. This implies that the novice group presented a higher level of mental workload that was reflected by fatigue (reflected by theta band power), awakening level (reflected by beta band power), memory (reflected by gamma band power) and attention (reflected by SMR band power).

Another study, titled "An evaluation of mental workload with frontal EEG", recorded the frontal EEG signals of 20 participants during four activities (arithmetic operation, finger tapping, mental rotation and a lexical decision task) in order to investigate dynamic changes in mental workload. The EEG output indicated that theta activity increased as the difficulty of tasks increased [17]. Mohanavelu et al. [18] used EEG to demonstrate the relationship between dynamic workload and two elements of cognitive workload and attention. A total of 16 male fighter pilots participated in the experiment. The researchers found that alpha band power and both high and low beta band powers, as recorded by the FT10, FP1, FC1, P4, P7, Pz, T8, CP2 and C4 sensors, were more dominant during the cruise phase of the study. In addition, the FC2, FP2, FT10, and C4 sensors indicated more significant levels of total beta band power during the landing phase in comparison with the other workload tasks.

Umar Saeed et al. [19] classified long-term stress with machine learning algorithms which utilized resting state EEG recording signals. They revealed that beta and gamma band powers, as measured by the AF3 sensor, were statistically significantly different in the stress and the control group (with a label assigned by expert evaluators used as the reference).

2.5 *Research Hypotheses*

The current study involves the evaluation of mental workload and stress during remote ship operations using EEG signals. Six hypotheses to assess the level of stress during remote ship operations are proposed.

Kari et al. [4] have identified high mental workload as a human factor issue which affects the performance of operators in SCCs. In SCCs, remote control systems should promote an optimal level of situational awareness by providing a high level of information, which increases the risk of high workload during remote ship operations. The impact of high workload as a primary human factor issue during remote operation has also been highlighted in previous studies [14, 20, 21]. Since we wanted to make sure that our experiments succeeded in manipulating the workload, the first hypothesis tests whether the level of workload was successfully manipulated during the experiments. The level of workload was assessed using the NASA task load index (TLX) to identify whether operators perceived a higher level of workload during the second scenario.

Remote operators can also experience higher levels of stress when they face more demanding tasks and higher mental workloads [11]. This indicates that there are connections between high mental workload and stress [11, 14]. Hence, the second hypothesis of this study is designed to investigate whether operators perceived a higher level of stress when a higher level of workload was imposed on them.

The Autonomous Ship Controller (ASC) sends a set of ship status indicators to the SCC. The SCC operators use these ship status indicators to monitor the overall status of the ship [22]. Two ship status indicators, weather and risk of collision, highly affect the mental workload and stress levels of operators [22, 1]. Van Buskirk et al. (2019) have proposed heavy weather ship handling simulation training to improve the competence of seafarers, because the need to make correct and time-sensitive ship handling decisions in heavy weather increases human stress levels and the risk of error [23]. In addition, Yoshida et al. (2021) have established that weather conditions, such as heavy rain and fog, increase the mental workload and stress levels of operators during autonomous surface ship operations, particularly in highly congested areas [24]. Hence, the third hypothesis is designed to assess the impact of harsh weather on stress levels during remote ship operations.

The human-machine interface (HMI) can greatly affect human performance during interactions with machines. Since SCC operators receive all their information from the HMI, the HMI's design may affect the human operators' performance during remote ship operations. Moreover, a well-designed HMI can facilitate access to processable situational information, which decreases the level of stress that operators are exposed to during remote ship operations [25]. In addition, there is a significant probability of human errors associated with HMI, and human errors associated with HMI will highly affect performance factors such as stress during the operation of autonomous ships [26]. Hence, the fourth hypothesis is designed to investigate the impact of HMI on the stress levels of operators during remote ship operations.

Radio communication (voice over VHF) is a standard method of communication between remote operators [14]. Distorted communication and background radio communication have been identified as two main factors which create stress for aerial firefighting pilots during training [21]. Moreover, levels of theta, alpha and beta EEG band powers in the posterior and left front-central areas of the brains of air control traffic operators seem to increase during stressful radio communication with airplane pilots [27]. During stressful radio communications, the number of clear speech events on the part of air control traffic operators is reduced, probably due to faster pronunciation [27]. Hence, the fifth hypothesis is designed to investigate how VHF radio communication impacts the stress levels of operators during remote ship operations.

Operators must be completely focused to avoid collision risks when investigating the vectors, status/heading and speed of the targets depicted by the collision indicator [22]. Perceived collision risks seem to increase the stress level of operators because of anxiety about collisions or the difficulty of performing collision avoidance navigation in close head-on or crossing situations [28]. Hence, the sixth hypothesis is designed to assess the impact of situations in which there is a risk of accidents on operators' stress levels.

In summary, this study will test the following hypotheses:

1 There is a significant change in the level of workload between the first and the second scenario in the experiments.
   - Corresponding null hypothesis: there is no significant change in the level of workload between the first and second scenarios in the experiments.
2 There is a significant change in stress when workload increases.
   - Corresponding null hypothesis: there is no significant change in stress when workload increases.
3 There is a significant change in stress when ships are operating in harsh weather.
   - Corresponding null hypothesis: there is no significant change in stress when ships are operating in harsh weather.
4 There is a significant change in stress when the number of ships increases.
   - Corresponding null hypothesis: there is no significant change in stress when the number of ships increases.
5 There is a significant change in stress when operators establish VHF communication.
   - Corresponding null hypothesis: there is no significant change in stress when operators establish VHF communication.
6 There is a significant change in stress when there is a risk of accident.
   - Corresponding null hypothesis: there is no significant change in stress when there is a risk of accident.

3 MATERIALS AND METHODS

In this study, a series of experiments was performed to evaluate the impact of workload and stress on operators of SSCs and thus to evaluate the proposed hypotheses.

3.1 *Instruments - EEG and NASA TLX*

Generally, workload and stress are measured subjectively by means of interviews or questionnaires. However, it is also possible to investigate changes in brain activity directly by using tools which measure biological processes. In this study, both direct and subjective measures were used. EEG was used for direct measurements and the NASA TLX system was used for subjective measurements. NASA TLX was mainly employed as a supportive technique to verify that manipulation of the workload, the independent variable, was successful and that participants were exposed to a higher workload in the second scenario.

EEG is used to record human brain signals, and our previous study showcased the applicability of EEG to the assessment of the stress levels of SCC operators under different workloads [29]. EEG records the electrical activity of the brain using electrodes, also called sensors. The electrodes are attached to the scalp to record the electrical potential generated by the brain [30]. Types of EEG systems differ according to the type of connection between the electrodes and the scalp surface; these types include dry and wet electrode EEG systems. Wet electrode EEG systems include gel, saline and semi-dry or water-based systems [30] and require the use of electrolytic liquid to improve conductivity. The EMOTIV EEG EPOC Flex saline kits which were utilized in this study are comprised of 32 electrodes. The EMOTIV EEG cap uses electrodes in the following locations: AFz (driven right leg), FCz (common mode sense), Fp1, Fp2, F7, F3, Fz, F4, F8, FT9, FC5, FC1, FC2, FC6, FT10, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, TP9, TP10, P7, P3, Pz, P4, P8, O1, Oz, and O2 [31]. The EEG EPOC FLEX passes signals through a few stages of processing. First, it processes data to remove sharp spikes, then passes data through a high-pass filter to remove the DV offset and slow drift. It then applies a Hanning filter before performing a fast Fourier transform (FFT). Band power is calculated from the square of the amplitude in each frequency bin and output is presented as uV^2 / Hz.

The NASA TLX system was developed by NASA Ames Research Center in the 1980s and is used to subjectively assess the workload of human operators working with human-machine interaction systems [32]. The NASA TLX is comprised of two instruments, a self-reporting questionnaire and comparison cards, and measures overall workload as the mean of weighted ratings. The self-reporting questionnaire is comprised of six questions, answered on a scale of 1-7, which are designed to assess levels of perceived workload and stress. The measurement of workload includes six subscales reflecting the independent variables mental workload, physical workload, temporal demand, frustration, effort and performance. The NASA TLX is based on an assumption that some combination of the aforementioned variables is likely to indicate the workload [33]. In the NASA TLX form, participants rate the performance questions from "perfect" to "failure", and other questions from "very low" to "very high" [34]. The comparison cards include the same six variables, and participants are asked to choose one item in each card.

3.2 *The experiments*

The experiments were performed in the navigation simulators of Norsk Maritim Kompetansesenter (NMK), a department of the Norwegian University of Science and Technology, Alesund, Norway. Three healthy male participants with no psychiatric problems or neurological disorders, participated in the experiments as SCC operators. The participants worked in the maritime domain but were not experts in the use of simulators. Before the experiments started, they were informed about the process and received written instructions for the experiments. In addition, informed consent was obtained from all subjects involved in the study. During the experiments, navigation simulators were used to represent an SCC (specifically the instructor room) and three ship bridge simulators were used to represent remotely controlled ships.

During the experiments, workload and stress were considered to be the independent and dependent variables respectively. On the basis of the status indicators in SCCs [35, 36, 22], the independent variable was manipulated by changing the number of targets (traffic), the number of ships to the SCC operator had to monitor, the difficulty of the route, the weather, and by introducing accident risks and establishing VHF communication between the SCC and ships. Table 1 illustrates the manipulation and measurement of the variables during the experiments.

Table 1. Types of workload and stress variables and how they were manipulated and measured.

Variables	Type of variable	Manipulation	Measurements
Workload	Independent	Number of targets, number of ships to be monitored by the SCC operators, difficulty of the route, weather, other events such as accidents	NASA-TLX technique Self-reporting questionnaire
Stress	Dependent		Physiological measurements of stress Raw EEG data Self-reporting questionnaire

Before the low and high workload scenarios, an initial scenario was performed to establish a baseline for the assessment of the impact of different levels of workload on brain activity, as well as for the identification of the trends and anomalies in the EEG signals. Figure 1 depicts a participant performing the baseline scenario while the EPOC FLEX was recording the EEG signals of his brain activity. In the baseline scenario, each participant sat in a comfortable chair in a calm and quiet environment and read a newspaper or book for 10–15 minutes.

Figure 1. A participant reading a book in a calm and quiet environment to establish baseline EEG brain activity.

The content of the low and high workload scenarios, which was discussed and approved in advance by three pilots (as experts in this domain), are presented in Table 2. Each of low and high workload scenarios were considered as a package of factors that may affect the level of workload perceived by remote ship operators.

Table 2. The high and low workload scenarios

Scenario	First (low workload)	Second (high workload)
Area	Kristiansund to Trondheim (low difficulty)	Vatlestraumen (moderate difficulty)
Number of ships	Three container ships (three-ship bridge simulators)	Five container ships
Traffic	5+ targets	15+ targets
Visibility	Good visibility in daylight	Bad visibility, nighttime
Weather	Moderated wind, calm sea-state	Strong wind, choppy sea
VHF communication	No	Yes
Risk of accident	No risk of accident	Two risks of accident
Overall workload	Low	High

Each experiment took 10–15 minutes due to the recording limitations of the EPOC Flex EEG. During the experiments, the EPOC Flex EEG recorded the brain activity of each participant via 32 sensors. Furthermore, a time recorder and a checklist were used to record events in order to synchronize the EEG data with external events. In addition, a video camera recorded activities in the SCC (instructor room) during the experiments to facilitate the correlation of external events with the operators' EEG signals. Each participant filled out the NASA TLX questionnaire and performed the comparison card exercise after each scenario in order to assess whether the workload increased in the second scenario and identify which factors were perceived by operators as demanding tasks during each experiment. In this way, the perception of high workload will be cross-validated by factors that operators perceived as demanding tasks during remote ship operations.. Figure 2 depicts a human operator performing the first scenario in the SCC, where the human operator was responsible for monitoring a ship.

Figure 2. A participant performing the low workload scenario (first scenario) in the SCC.

To simulate the monitoring mode of SCCs, during the experiments participants were responsible for monitoring the status and route of each ship and, if necessary, sending high-level commands to the ship. The participants monitored ships' status indicators, including speed, rate of turn, heading, engine status, rudder status, and propeller revolution. In cases of red alarms, participants were responsible for informing the ships via VHF communication.

Since the experiments involved three scenarios for three participants, nine sets of EEG data and NASA TLX self-reporting questionnaires and rating cards were produced. The scores of the rating sheet and rating cards were analyzed to calculate the overall workload.

3.3 *Analysis*

The EEG signals were analyzed by SPSS and a cloud-based visualization platform (Kibana). The EEG dataset comprised 160 features and a total of 42,084 samples, because signal of each EEG sensor preprocessed to generate five band powers including alpha, low beta, high beta, theta and gamma. Samples in the dataset were thus labeled with a binary value for the workload variable (where 0 = low workload and 1 = high workload). SPSS was used to calculate the Pearson correlation coefficient matrix and a correlation coefficient for each of 160 band powers. The Pearson correlation coefficient matrix was then used to identify which EEG band powers correlated with changes in workload and stress. A cloud-based visualization platform using Elastic Stack [37] was used to analyze the EEG data and identify trends and anomalies. Finally, the EEG data were correlated with workload variables to identify how the brain activity of human operators changes under changes in workload and stress.

The NASA TLX system analysis a two-part evaluation process comprised of rating and weighting processes. There were 15 pair-wise comparison cards for the six scales. On each card, participants circled the member of each pair that contributed more to the workload. In addition, participants filled out the rating sheet with a numerical rating for each scale. The overall workload score for each participant was calculated by multiplying each rating by the relevant weighting factor. Finally, the sum of the weighted ratings was divided by 15 (15 being the sum of the weights) [38].

4 RESULTS

The results of the NASA TLX analysis are presented graphically in Figure 3 to distinguish the overall workloads perceived by each participant after each scenario. Based on the NASA TLX technique, the overall workloads of the first participant were calculated as 3.2 and 18.5 during the low and high workload scenarios respectively. The overall workloads of the second participant were calculated as 7.5 and 11.86 during the low and high workload scenarios respectively. The overall workloads of the third participant were calculated as 5.2 and 16.6 during the low and high workload scenarios respectively. Figure 3 depicts the calculated perceived overall workload of each participant in the experiments. As can be seen in Figure 3, all participants perceived a higher level of workload during the second scenario.

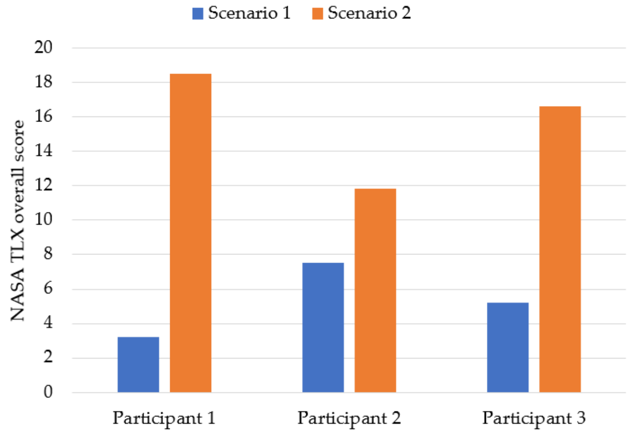

Figure 3. Calculated perceived overall workload of each participant in the first and second scenarios.

The results of a paired samples t-test, including the mean difference, t-value and two-tailed probability of each variable, are presented in Table 3. According to the sampling distribution of t, the t-value was 4.303 for the two-degree field for the rejection of a null hypothesis, with a 95% confidence interval (CI) and 0.05 significance level. Furthermore, the 0.199 p-value was greater than the 0.05 alpha level, indicating that there was no significant change in overall stress between the baseline and low workload scenarios, with a 95% CI of mean difference [-5.46, 2.12].

Table 3. Statistical analysis of self-reporting questionnaires

Variable	Mean. Diff	t-value	Sig. (2-tailed)
Overall stress (baseline-low workload)	-1.666	-1.89	0.199
Overall stress (baseline-high workload)	-5.00	-8.66	0.013
VHF communication	-2.66	-3.02	0.94
Risk of accident	-1.333	-1.51	0.27
Weather	-3.66	0.11	0.008
Number of ships	-4.00	-6.92	0.02
Overall workload	-2.66	-8.00	0.15

The results indicate that there was a significant difference between the baseline and high workload scenarios: the 95% confidence interval [-7.48, -2.51] did not contain zero. In addition, the p-value was lower than the 0.05 alpha level, which also indicates there was a significant difference between baseline and high workload scenarios.

The participants did not report higher levels of stress when establishing VHF communication or when there was an increased risk of accident in the high workload scenario. For VHF communication, a 0.94 p-value that was greater than the 0.05 alpha level indicated that there was no significant difference in stress between the low and high workload scenarios during VHF communication. In addition, the t-value (3.024) was less than the critical t-value, which also indicated there was no significant difference in stress between low and high workload scenarios during VHF communication. For increased risk of accidents, a p-value (0.27) greater than the 0.05 alpha level indicated there was no significant difference in stress between low and high workload scenarios when there was a risk of accident. In addition, the t-value (-1.51) was less than the critical t-value, which also indicated there was no significant difference in stress between low and high workload scenarios when there was a risk of accident.

Participants reported higher levels of stress in wavy waters and harsh weather and when the number of ships was increased. For weather, the p-value (0.008), t(2) (11.00) and the 95% CI [-5.10, -2.23] all indicated that there was a significant difference in stress between the low and high workload scenarios. In addition, the p-value was less than the 0.05 alpha level, which also indicated there was a significant difference in stress when the water was wavy and the weather was harsh. For the increase in the number of ships, the p-value (0.02) was less than the 0.05 alpha level, indicating that there was a significant difference in stress between the low and high workload scenarios. In addition, the size effect (0.86) indicated that there was a positive correlation between the increase in the number of ships and the participants' stress levels.

Participants reported higher levels of overall mental workload during the high workload scenario. The statistical analysis indicated that there was a significant change in the overall mental workload during the high workload scenario: the p-value was 0.15, the t(2) was -8.00, and the 95% CI [-4.10, -1.23] did not contain zero.

Thus, based on the statistical analysis of the NASA TLX self-reporting questionnaires and comparison cards, it can be concluded that two variables—harsh weather and the number of ships—affected the workload, and consequently the stress levels, of the SCC operators in the experiments.

The samples in the EEG dataset were labeled with the corresponding values of the manipulated factors (weather, number of ships, risk of accident, etc). In this study, EEG band powers were considered as dependent variables, while the manipulated factors were considered as independent variables. According to the correlation coefficient matrix, two EEG band powers—gamma and beta—had the highest correlation with the independent variables. This indicates that gamma and beta band powers significantly increased when the number of ships that the participants had to monitor increased. This study follows the Pearson correlation coefficient classification: high ($\pm 0.50 \leqslant$ high $\leqslant \pm 1$), moderate ($\pm 0.30 \leqslant$ moderate $< \pm 0.50$) and low correlation ($\pm 0.1 <$ low $\leqslant \pm 0.29$). Figure 4 depicts the EEG sensors with high (purple), moderate (green) and low (blue) correlations with the weather and ship number variables in a 10–20 EEG sensor placement system. Figure 4.a illustrates that EEG sensors for the first participant had low correlation with the number of ships and with weather status. In Figure 4.a, sensors with low correlation to these two variables, including F3, FC1, TP9, TP10, P4, O1, Oz and O2, are colored in blue. Figure 4.b depicts the EEG sensors for the second participant which had moderate and high correlations with the number of ships and with weather status. Sensors with high correlation, including F7 and T8, are colored in purple, while sensors with moderate correlation, including Fp1, FP2, F4, F8, TP9, TP10 and P7, are colored in green. Figure 4.c depicts the EEG sensors for the third participant which had moderate and high correlations with the number of ships and with weather status. In

Figure 4.c the FT9 sensor with moderate correlation is colored in green, while the FT10 sensor with high correlation is colored in purple.

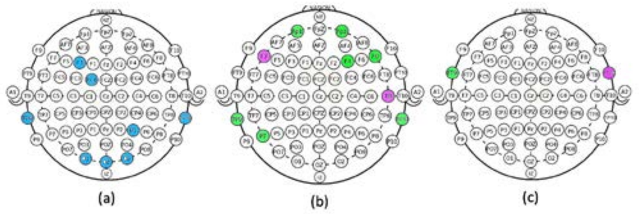

Figure 4. EEG sensors that indicated high (purple), moderate (green) or low (blue) correlations with increases in workload and stress; (a) denotes participant 1, (b) denotes participant 2, and (c) denotes participant 3.

Figure 5 depicts the EEG signals of beta and gamma band powers recorded during the experiments, where the first, second and third graphs illustrate baseline, low workload and high workload scenarios respectively in each sub-figure. The levels of EEG measurements were different in each scenario where sensors with moderate and high correlations presented considerable brain activity changes than sensors with low correlation. Hence, Figure 5 depicts brain activity levels in each scenario for the low (participant 1), moderate (participant 2) and high (participant 3) correlation group of sensors. Figure 5 illustrates the level of changes for low, moderate and high correlation sensors therefore sensors were selected randomly for demonstration of brain activity changes during baseline, low and high workload scenarios. While the calculated correlation of all EEG sensors of participant 1 were low thus Figures 6.e and 6.f depicts brain activity changes of low correlation sensors. Because sensors of participant 2 presented moderate and high correlations, Figure 5.c depicts brain activity measured by a sensor with moderate correlation while Figure 5.d depicts brain activity measured by a sensor with high correlation for participant 2. To show the changes of brain activity measured by different band powers, Figures 6.a and 6.b depict brain activity measured by different band powers of a sensor with high correlation for participant 3. Figure 5.a indicates the EEG signal of gamma band power of the FT10 sensor for the third participant. As can be seen in Figure 5.a, the level of gamma band power significantly increased when the workload increased in the high workload scenario. Figure 5.b indicates the EEG signal of beta band power of the FT10 sensor for the third participant. Figure 5.b shows that levels of both beta and gamma band powers significantly increased when the workload increased in the high workload scenario. Figure 5.c depicts the EEG signal of the gamma band power of the P7 sensor for the second participant. Figure 5.c shows that the level of gamma band power also significantly increased in the high workload scenario. Figure 5.d depicts the EEG signal of the gamma band power of the T8 sensor for the second participant. Figure 5.d shows that the level of gamma band power also significantly increased when the workload increased in the high workload scenarios. Figure 5.e depicts the EEG signal of the gamma band power of the FC1 sensor for the first participant. Figure 5.e shows that the level of gamma band power changed slightly between the baseline, low workload and high workload scenarios. Figure 5.f depicts the EEG signal of the gamma band power of the F3 sensor for the first participant. Figure 5.f shows that the level of gamma band power also changed slightly between the low and high workload scenarios.

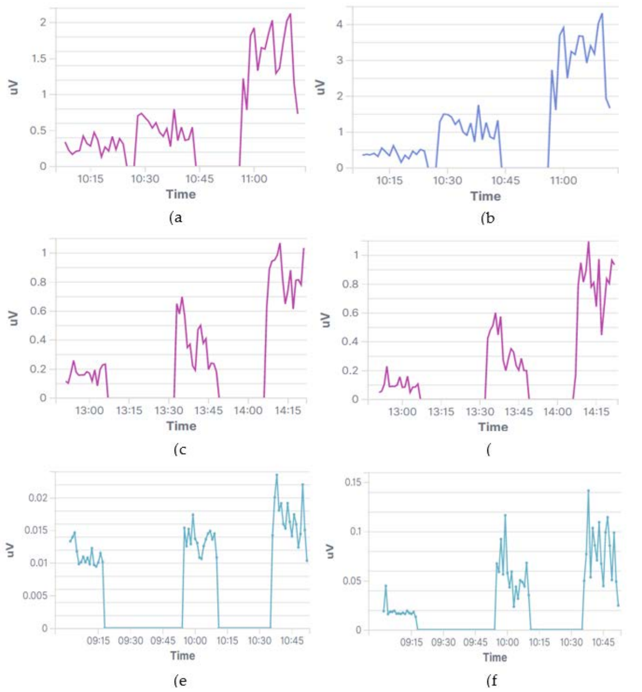

Figure 5. Visualization of EEG band powers in uV during the baseline, low workload and high workload scenarios: (a) beta band power of FT10 sensor for participant 3; (b) gamma band power of FT10 sensor for participant 3; (c) gamma band power of P7 sensor for participant 2; (d) gamma band power of T8 sensor for participant 2; (e) gamma band power of FC1 sensor for participant 1 ; and (f) gamma band power of F3 sensor for participant 1

5 DISCUSSION

This study investigated human factor challenges during remote ship operations and highlighted the different human factors involved. It is evident that one of the main challenges is an increase in the mental workload of SCC operators due to operational tasks. SCC designers aim to identify the maximum workload level for the efficient performance of remote operations by SCC operators.

The current study focuses on variables that may increase the level of mental workload of SCC operators, such as the number of ships that they are responsible for, traffic, weather conditions, VHF communication and the risk of accidents. The correlation matrix of the EEG results indicates that the gamma and beta band powers of the FT10, P7 and T8 sensors were highly correlated with weather status and the number of ships to be monitored. The gamma and beta band powers were, in fact, the only band powers that recorded changes in workload and stress levels in all participants. The results from the statistical analysis of the self-reported NASA TLX data also indicate significant changes in stress levels when ships are operating in harsh weather and when the number of ships is increased. When the number of ships were increased, number of human machine interfaces (HMIs) that an operator should interact during experiments increased considerably. The way that operators received information from HMI also affected the level of stress because operators should collect critical information in a short time span for more than one ship. In addition, significant increase of P7 sensor (please see Figure 5c) which covers inferior lateral occipital cortex responsible for eye movements regarding object recognition in a visual information collection process supports the impact of HMI on stress when the number of ships increases. Furthermore, either low or no changes in stress were recorded when operators established VHF communication or when there was a risk of accidents. The direct measurement of brain activity by EEG and the subjective self-reported findings therefore support each other with regard to hypotheses 3, 4, 5 and 6, which make the findings more credible.

All participants perceived a higher mental workload during the high workload scenario. Hence, this study successfully managed to manipulate mental workloads in the low workload and the high workload scenarios, which supports hypothesis 1. Since overall stress and workload increased during the high workload scenario, hypothesis 2 is also supported. Increase in the number of ships the operators were responsible for and worsening of the weather both had significant impacts on stress levels, and therefore hypotheses 3 and 4 are also supported. The results show, however, that establishing VHF communication and increasing the risk of accidents did not have significant impacts on operators' stress levels, and therefore hypotheses 5 and 6 are not supported. Hence, four hypotheses (1, 2, 3 and 4) were accepted, while two hypotheses (5 and 6) were not accepted. Support for each hypothesis according to the experimental results is summarized in Table 4.

Table 4. Hypotheses test results

H#	Hypothesis	Result
1	There is a significant change in the level of workload between the first and the second scenario in the experiments	Supported
2	There is a significant change in stress when workload increases	Supported
3	There is a significant change in stress when ships are operating in harsh weather	Supported
4	There is a significant change in stress when the number of ships increases	Supported
5	There is a significant change in stress when operators establish VHF communication	Not supported
6	There is a significant change in stress when there is a risk of accident	Not supported

This study also has some limitations. The number of participants is low, and the participants are not experienced SCC operators.

6 CONCLUSION

This study performed human-centered experiments to investigate the stress levels of SCC operators during human-human and human-machine interactions, and tested six hypotheses to assess the human factors of workload and stress. Nine experiments were performed to collect the brain activity of human operators using EEG equipment, resulting in a dataset consisting of more than 42,000 samples. In addition, the NASA TLX test was used so that the operators could self-assess workload and stress levels. On the basis of the statistical analysis, four hypotheses were accepted while two were rejected. In addition, a correlation coefficient matrix was generated to identify correlations between the brain activity of operators and workload and stress levels. This indicated that the beta and gamma band powers of the EEG recordings were highly correlated with workload and stress levels during remote ship operations. The results show that increases in workload result in significant changes in stress levels when ships are operating in harsh weather and when the number of ships each SCC operator is responsible for increases. The results also show that there is no significant change in stress levels when SCC operators establish VHF communication or when there is a risk of accidents. The practical implications of these findings are that SCC designers, SCC operator training programs and standardization bodies can utilize these results to improve the safety and efficiency of remote ship operations.

Future studies should investigate other human factors affecting workload and stress levels in remote ship operations. Future studies are also needed to perform these experiments with experienced SCC operators in order to improve the applicability of the results of this study. Moreover, studies with more participants are needed. It would also be interesting to extend this study by performing machine learning processes on EEG signals to provide a platform for customizing operator training programs and improving SCC designs and protocols.

REFERENCES

1. Rødseth ØJ, Faivre J, Hjørungnes SR, Andersen P, Bolbot V, Pauwelyn A-S, Wennersberg LA (2020) AUTOSHIP deliverable D3.1 Autonomous ship design standards, Revision 2.0.
2. Rødseth, Ørnulf & Tjora, Åsmund. (2014). A system architecture for an unmanned ship.
3. Kim, M., Joung, T. H., Jeong, B., & Park, H. S. (2020). Autonomous shipping and its impact on regulations, technologies, and industries. Journal of International Maritime Safety, Environmental Affairs, and Shipping, 4(2), 17-25.
4. Kari, R.; Steinert, M. Human Factor Issues in Remote Ship Operations: Lesson Learned by Studying Different Domains. J. Mar. Sci. Eng. 2021, 9, 385. https://doi.org/10.3390/jmse9040385
5. MO MSC, 2021. Regulatory Scoping Exercise for the Use of Maritime Autonomous Surface Ships (MASS) (No. 99/WP.9). London.
6. S. N. MacKinnon, Y. Man, and M. Baldauf, "D8.8: Final Report: Shore Control Centre." Maritime Unmanned Navigation through Intelligence in Networks, 2015.
7. Man, Y., Weber, R., Cimbritz, J., Lundh, M., & MacKinnon, S. N. (2018). Human factor issues during remote ship monitoring tasks: An
8. Rødseth, Ørnulf & Nordahl, Håvard. (2018). Definition of autonomy levels for merchant ships, Report from NFAS, Norwegian Forum for Autonomous Ships, 2017-08-04.. 10.13140/RG.2.2.21069.08163.
9. Kim, M., Joung, T. H., Jeong, B., & Park, H. S. (2020). Autonomous shipping and its impact on regulations,

technologies, and industries. Journal of International Maritime Safety, Environmental Affairs, and Shipping, 4(2), 17-25.

10. Zhu, T., Haugen, S., & Liu, Y. (2019, September). Human factor challenges and possible solutions for the operation of highly autonomous ships. In Proceedings of the 29th European Safety and Reliability Conference, Hannover, Germany (pp. 22-26).
11. Grech, M.R., Horberry, T., & Koester, T. (2008). Human Factors in the Maritime Domain.
12. Wahlström, M.; Hakulinen, J.; Karvonen, H.; Lindborg, I. Human factors challenges in unmanned ship operations-insights from other domains. Procedia Manuf. 2015, 3, 1038–1045.
13. Man, Y., Lundh, M., Porathe, T., & MacKinnon, S. (2015). From desk to field-Human factor issues in remote monitoring and controlling of autonomous unmanned vessels. Procedia Manufacturing, 3, 2674-2681.
14. Alsuraykh, N. H., Wilson, M. L., Tennent, P., & Sharples, S. (2019, May). How stress and mental workload are connected. In Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare (pp. 371-376)
15. Dussault C, Jouanin J-C, Philippe M, Guezennec C-Y. EEG and ECG changes during simulator operation reflect mental workload and vigilance. Aviat Space Environ Med 2005; 76:344–351.
16. Ma, Qing & Shang, Qian & Fu, Hui & Chen, Fu. (2012). Mental Workload Analysis during the Production Process: EEG and GSR Activity. Applied Mechanics and Materials. 220-223. 193-197. 10.4028/www.scientific.net/AMM.220-223.193.
17. So, W. K., Wong, S. W., Mak, J. N., & Chan, R. H. (2017). An evaluation of mental workload with frontal EEG. PloS one, 12(4), e0174949
18. Mohanavelu, K., Poonguzhali, S., Adalarasu, K., Ravi, D., Chinnadurai, V., Vinutha, S., ... & Jayaraman, S. (2020). Dynamic cognitive workload assessment for fighter pilots in simulated fighter aircraft environment using EEG. Biomedical Signal Processing and Control, 61, 102018.
19. Saeed, S. M. U., Anwar, S. M., Khalid, H., Majid, M., & Bagci, U. (2020). EEG based classification of long-term stress using psychological labeling. Sensors, 20(7), 1886.
20. Gjoreski, M., Luštrek, M., Gams, M., & Gjoreski, H. (2017). Monitoring stress with a wrist device using context. Journal of biomedical informatics, 73, 159-170.
21. Clifford, R.M.S., Engelbrecht, H., Jung, S. et al. Aerial firefighter radio communication performance in a virtual training system: radio communication disruptions simulated in VR for Air Attack Supervision. Vis Comput 37, 63–76 (2021). https://doi-org.ezproxy.uio.no/10.1007/s00371-020-01816-6
22. Ø. J. Rødseth, B. Kvamstad, T. Porathe and H. -C. Burmeister, "Communication architecture for an unmanned merchant ship," 2013 MTS/IEEE OCEANS - Bergen, 2013, pp. 1-9, doi: 10.1109/OCEANS-Bergen.2013.6608075.
23. Van Buskirk L.J., Alman P.R., McTigue J.J. (2019) Further Perspectives on Operator Guidance and Training for Heavy Weather Ship Handling. In: Belenky V., Spyrou K., van Walree F., Almeida Santos Neves M., Umeda N. (eds) Contemporary Ideas on Ship Stability. Fluid Mechanics and Its Applications, vol 119. Springer, Cham. https://doi-org.ezproxy.uio.no/10.1007/978-3-030-00516-0_49
24. Yoshida, M.; Shimizu, E.; Sugomori, M.; Umeda, A. (2021) Identification of the Relationship between Maritime Autonomous Surface Ships and the Operator’s Mental Workload. Appl. Sci. 2021, 11, 2331. https://doi.org/10.3390/app11052331
25. Kimberly Tam, Rory Hopcraft, Tom Crichton & Kevin Jones (2021) The potential mental health effects of remote control in an autonomous maritime world, Journal of International Maritime Safety, Environmental Affairs, and Shipping, 5:2, 40-55, DOI: 10.1080/25725084.2021.1922148
26. Liu, J., Aydin, M., Akyuz, E. et al. Prediction of human–machine interface (HMI) operational errors for maritime autonomous surface ships (MASS). J Mar Sci Technol (2021). https://doi-org.ezproxy.uio.no/10.1007/s00773-021-00834-w
27.[G. Borghini et al., "Stress Assessment by Combining Neurophysiological Signals and Radio Communications of Air Traffic Controllers," 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2020, pp. 851-854, doi: 10.1109/EMBC44109.2020.9175958.
28. Kim, D. H. (2020). Human factors influencing the ship operator's perceived risk in the last moment of collision encounter. Reliability Engineering & System Safety, 203, 107078. work. International Journal of Industrial Ergonomics, 86, 103233.
29. Kari, R., Steinert, M., & Gaspar, H. M. (2019). Eeg application for human-centered experiments in remote ship operations. In CENTRIC 2019, The Twelfth International Conference on Advances in Human oriented and Personalized Mechanisms, Technologies, and Services. International Academy, Research and Industry Association (IARIA)
30. BitBrain, The Wet EEG Cap & Differences Between Water-Based, Saline and Gel EEG caps. Available online: https://www.bitbrain.com/blog/wet-eeg-cap (accessed on 16 January 2022)
31. Emotive, EEG EPOC FLEX. Available online: https://emotive.com (accessed on 16 January 2022)
32. NASA TLX: Task Load Index. Available online: https://humansystems.arc.nasa.gov/groups/TLX/ (accessed on 16 January 2022)
33. Hart, S. G. (2006, October). NASA-task load index (NASA-TLX); 20 years later. In Proceedings of the human factors and ergonomics society annual meeting (Vol. 50, No. 9, pp. 904-908). Sage CA: Los Angeles, CA: Sage publications.
34. Braarud, P. Ø. (2021). Investigating the validity of subjective workload rating (NASA TLX) and subjective situation awareness rating (SART) for cognitively complex human–machine
35. Vindøy, V. (2008). A functionally oriented vessel data model used as basis for classification. 7th International Conference on Computer and IT Applications in the Maritime Industries, COMPIT, 8.
36. Rødseth, Ørnulf Jan, Drezet, F., Pedersen, E. S., Jensen, N. A., Ehrke, K.-C., Oma, P. N., & Giere, R. (2008). TCI and status indicator specification, Flagship deliverable D-D1 - Google Search. Retrieved May 29, 2019, from https://www.google.com/search?client=safari&rls=en&q=TCI +and+status+indicator+specification,+Flagship+deliverable+ D-D1&ie=UTF-8&oe=UTF-8
37. Elastic. Elastic Stack, Elastic. Available online: https://www.elastic.co/elastic-stack (accessed on 19 January 202
38. NASA TASK LOAD INDEX (TLX) v. 1.0. Paper and Pencil Package. Human Performance Research Group. NASA Ames Research Center. Moffett Field, California.
39. Seo, S.-H., & Lee, J.-T. (2010). Stress and EEG. In Convergence and hybrid information technologies. IntechOpen.
40. Harmony, T., Fernández, T., Silva, J., Bernal, J., Díaz-Comas, L., Reyes, A., ... Rodríguez, M. (1996). EEG delta activity: An indicator of attention to internal processing during performance of mental tasks. International Journal of Psychophysiology, 24(1–2), 161–171
41. Rajendran, V. G., Jayalalitha, S., & Adalarasu, K. (2021). EEG Based Evaluation of Examination Stress and Test Anxiety Among College Students. IRBM.
42. Luijcks, R., Vossen, C. J., Hermens, H. J., van Os, J., & Lousberg, R. (2015). The Influence of Perceived Stress on

Cortical Reactivity: A Proof-Of-Principle Study. PloS one, 10(6), e0129220. https://doi.org/10.1371/journal.pone.0129220
43. Miskovic V, Ashbaugh AR, Santesso DL, McCabe RE, Antony MM, Schmidt LA. Frontal brain oscillations and social anxiety: a cross-frequency spectral analysis during baseline and speech anticipation. Biol Psychol. 2010 Feb;83(2):125-32. doi: 10.1016/j.biopsycho.2009.11.010. Epub 2009 Nov 27. PMID: 19945500