AN INVESTIGATION OF THE INFLUENCE OF A DECISION SUPPORT SYSTEM ON SIMULATED ICE MANAGEMENT PERFORMANCE

By Jonathan Soper

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of

Master of Engineering

Department of Ocean and Naval Architectural Engineering

Faculty of Engineering and Applied Science

Memorial University of Newfoundland and Labrador

October 2022

St. John’s, Newfoundland & Labrador

Abstract

As automation becomes increasingly widespread in the marine industry, it is important that the efficacy of automated systems be appropriately gauged and measured. In this research, an experimental method was applied to test a decision support system (DSS) using inexperienced participants. These participants are provided with basic habituation but otherwise must depend on the DSS to provide strategic advice. They are then tasked with using their vessel to clear pack ice from the lifeboat launch zone of an offshore structure, using the DSS for strategic assistance. The experimental simulator has been used in other studies to gauge the effects of experience and training on ice management performance, which allows for a comprehensive comparison to be made across several different groups. The DSS contains expert knowledge and acts as a digital coach for the inexperienced participants in this study. It provides advice through video replays and text-based advice that was customized to their current approach when assistance is requested. The participants tended to adopt the expert strategies suggested by the DSS, but failed to deliver statistically significant improvements. This can be in part attributed to their lack of ship handling experience. Qualitative analysis from exit interviews also provided a good basis to guide improvements to the DSS. Ultimately, the experiment provides insights into the strengths and weaknesses of this DSS and describes an effective experimental method for evaluating this technology.

Acknowledgements

I would like to gratefully acknowledge the National Research Council of Canada (NRC) Ocean Program for the financial sponsorship of this research. I would also like to thank the Natural Sciences and Engineering Research Council (NSERC)/ Husky Energy (Cenovus) Industrial Research Chair in Safety at Sea for their financial support. Thank you to Virtual Marine Technologies for their technical support and assistance with the simulator throughout the experimental campaign.

A considerable contribution to the collection and analysis of data was made by the ONAE work term student, Patrick Maloney, who helped to troubleshoot and run the simulator, and work with our various analysis scripts. Thank you for all your hard work in your term with us.

I'd also like to thank my supervisors- Dr. Brian Veitch for providing me with the opportunity to work in his lab, Dr. Jennifer Smith for her experience and guidance with experimental design, the workings of the simulator, and scientific writing, and Dr. Jon Power for his contributions to the analysis of my results and editing of my writing. The guidance of all three supervisors during our biweekly meetings was extremely beneficial to keeping me focused and on track.

My passion for the marine industry would not have been fostered without the guidance, support, and opportunities provided by the Tall Ship St. Lawrence II out of Kingston, Ontario, and my team at Kingston 1000 Islands Cruises. Thank you to captains Dugald Henderson, Christopher Chafe, Norman Rennie, Mel Boyd, and Stephen Steels for guiding me over many years into this career path and encouraging me to achieve various Mate and Master's tickets, and trusting me to safely command their vessels.

Finally, I would like to thank my family, parents Tim and Wendy, and sister Alice for always encouraging a pursuit of learning, building, and sailing. That mindset growing up helped me to pursue a career in engineering. And thank you to my partner Kate Charlesworth for helping to make St. John's home and joining me on this adventure

Contents

List of Tables

List of Figures

List of Equations

List of Abbreviations and Symbols

AHTS	Anchor Handling Tug Supply
ANOVA	Analysis of Variance
CRD	Completely Random Design
DSS	Decision Support System
FPSO	Floating Production, Storage, and Offloading
IMO	International Maritime Organization
ECDIS	Electronic Chart Display and Information System
STCW	Standards of Training, Certification and Watchkeeping
μ	Mean
σ	Standard Deviation

.

1 Introduction

1.1 Overview

As improvements in technology allow for the implementation of autonomous navigation systems in the marine industry, it is important to test the effects of these systems on marine navigational performance by human users. In this research, the efficacy of an autonomous decision support system (DSS) that was built to assist seafarers in maritime ice management operations is tested using a marine simulator.

Automation is often described as a way to reduce the dependence on human intervention in tasks with the expectation of reducing accidents overall. As automation becomes more widespread in the maritime field, it is important to ensure that any technological solutions are not over-promised or expected to provide unrealistic benefits. Further, it is important to ensure that best practices of human-centred design are incorporated into any new technology to avoid any negative unintended consequences caused by said technology. It is often unfairly reported that human error is the largest cause of marine accidents (Wróbel, 2021). These reports often do not take into account that technological changes have the possibility to introduce new factors that may confuse operators and contribute themselves to an accident, as was seen in the high-profile accident of the U.S.S. John S. McCain (National Transportation Safety Board, 2017). Since technological advances could confound conclusions of safety, it is important to ensure that any new automated devices are designed in such a way that they do not degrade the situational awareness of seafarers (Mallam, et al., 2020).

Ice management is performed in many jurisdictions around the world, from Bohai Bay to the Baltic Sea. In Canada, each winter and spring brings new challenges for marine shipping as regions in the Great Lakes, Gulf of St. Lawrence, coastal Newfoundland and Labrador, and the Arctic Sea all experience sea ice conditions (Dunderdale & Wright, 2005). The experience of the seafarers working in these conditions is often highly variable and non-standardized, as there are large differences in the conditions a seafarer might experience over their career. The Canadian Coast Guard College has begun to design and implement simulator training courses for ice management, however the specific tasks of a Coast Guard vessel generally differ from those of the Canadian offshore oil and merchant shipping industries (Canadian Coast Guard College, 2018). Training which is provided in the oil industry is often on-the-job training, and regulatory requirements are generally based on arctic regulations that don't directly translate to pack ice management in the offshore (Smith, et al., 2020). This is particularly true for seafarers working on offshore supply vessels in the offshore oil industry of the Newfoundland Grand Banks. These vessels are tasked to manage pack ice so that it does not cause risk to the safety of personnel and the environment or impede offshore operations. For example, if pack ice were to surround an offshore installation such as a floating oil drilling platform, there is the potential for the platform to be displaced farther than its allowable limits. Production must be halted to avoid damage to the drilling equipment, resulting in costly and potentially environmentally damaging consequences. Additionally, typical lifeboats on offshore installations require open water to launch. If pack ice enters the lifeboat's launch zone, this could impede an evacuation should an emergency occur (Simões Ré & Veitch, 2007). Variable ice conditions year to year can lead to uneven levels of crew experience. This has the potential to lead to variable outcomes in ice management performance.

Introducing a low-level form of automation such as a DSS to the bridge of a ship could be a useful way of transferring expert information to less experienced seafarers and in turn, reducing the variability and increasing the standard of ice management performance. A DSS is considered a Degree One form of automation by the International Maritime Organization (IMO), meaning it provides the user with decision support and guidance but still leaves all critical decision making to them (International Maritime Organization, 2021). This is the most basic out of four levels, with Degree Four being full autonomy. This work aims to test such an early stage DSS by evaluating the ice management performance of inexperienced individuals in a simulated environment. This DSS was first developed by Yazdanpanah (2021), and was updated for this experiment. The DSS utilizes knowledge captured from expert seafarers and delivers it to new, less experienced users in an attempt to improve their ice management performance (Smith, Yazdanpanah, Thistle, Musharraf, & Veitch, 2020). Ultimately, the performance improvement effects of the DSS factor will be tested against factors of experience and training. Datasets for experience and training were acquired through the experimental campaigns of Veitch et al. (2019) and Thistle & Veitch (2020) respectively, providing the current experiment with a strong background to which comparisons can be made.

1.2 Background

Previous experiments to gauge the effects of experience and training on ice management performance were done using an ice management research simulator (Veitch et al., 2018, Thistle &Veitch., 2020). In the current research, an experiment was designed to formally test the effects of an early-stage DSS system on ice management against the aforementioned factors. The first experiment, which will be referred to as Experiment I, was performed by Veitch et al. (2018), and tested the effects of seafarer experience. Two groups were selected to complete a 2^2 full-factorial experiment (Montgomery, 2013). The low-level experience group was chosen from inexperienced nautical science students, or cadets, from a local college program. None of these participants were provided with any training in the simulation. The high-level experience group was drawn from Canadian seafarers who had an average of 20 ± 10 years of experience working at sea (Veitch, Molyneux, Smith, & Veitch, 2018). All of the seafarers had experience working in ice, whether it was working in Newfoundland's offshore industry, to icebreaking in the Arctic. In order to ensure that the effect on ice management performance was being gauged independently, low ($4/10^{th}$) and high ($7/10^{th}$) levels of ice concentration were also tested, which made it a two-factorial experiment (Veitch, Molyneux, Smith, & Veitch, 2018).

A second experiment examined the effects of training on inexperienced participants. This experiment, known as Experiment II, by Thistle and Veitch (2019) once again made use of the research simulator to test the effect of training on ice management activity. For this experiment, cadets were recruited and provided with either one, or two sessions of ice management training in the simulator, corresponding to low and high levels respectively (Thistle & Veitch, 2019). They were also split into low and high level ice concentration groups. These results were analyzed against cadets and seafarers from Experiment I to measure the effects of training.

In the thesis work by Yazdanpanah (2021), the results from Experiments I and II were compiled into a case-based reasoning (CBR) DSS. This DSS would allow a new user in the simulator to request assistance, and have their current simulated situation cross-referenced with the cases from Experiments I and II in a database of cases, known as a case base. The closest approach would be matched and presented to users. Instructions were also provided based on interviews with other experienced seafarers who critiqued the previous cases, and gave their best practices for ice management. This DSS was used as the basis for the current experiment, which will be referred to as Experiment III.

In Experiment III, the DSS from Yazdanpanah (2021) was tested in a pilot study, and adjustments were made based on these results in an attempt to improve the system. Once the DSS design was finalized, inexperienced participants were recruited to test the effects of a digital DSS on ice management performance for inexperienced seafarers. 12 nautical science cadets, and 6 post-secondary students in maritime-related programs were recruited as they were considered inexperienced. They were then tested in the same simulator as used in Experiments I and II so that comparisons could be made across all groups.

1.3 Purpose

The purpose of this research was to test the effects of a digital DSS on the performance of inexperienced marine operators when performing an ice management task in the simulator. The experiment was designed to compare the results of the inexperienced groups guided by the DSS to the findings in Experiments I and II, which investigated the effects of experience and training, respectively on ice management performance. Additionally, the findings of exit interviews with participants were used to determine what improvements could be made to similar technology in future development.

1.4 Hypotheses

H_1: It was hypothesized that the group using the DSS, entitled the DSS Cadets, who are the participants in the current experiment, would perform the ice management task significantly better than the No DSS Cadets, who are the untrained, inexperienced participants in Experiment I. It was also hypothesized that the DSS Cadets' performance would be somewhere in between that of the No DSS Cadets, and the seafaring participants from Experiment I. Additionally, it was hypothesized that the DSS cadets would be between that of the No DSS cadets and the Training I and Training II participants from Experiment II. In that experiment, it was found that increasing training levels had a tendency to reduce variability and significantly improve performance (Thistle & Veitch, 2019). It was thought that the DSS would increase performance in a similar way. These trends were expected to be reflected in all performance metrics evaluated. The reasoning behind these hypotheses was that the DSS would assist the cadets by delivering expert strategic advice that would enable them to perform in a manner more closely resembling that of experienced seafarers or trained cadets. An example of the expected hypothesis can be seen in Figure 1. These results were generated manually and are for a generalized performance metric to illustrate the hypothesis.

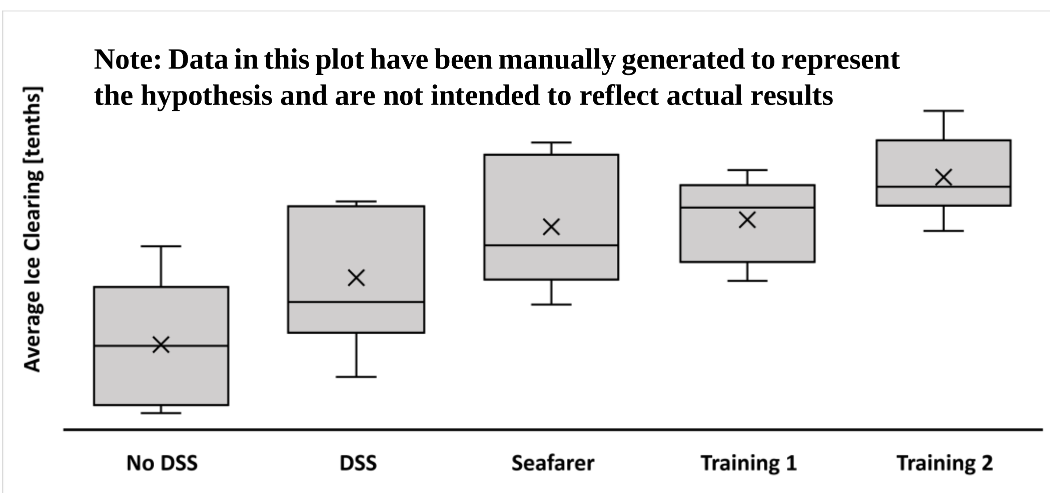


Figure 1: Manually generated hypothesized results for a generalized performance metric.

H_2: It was hypothesized that the ice concentration would not result in significant differences among the performance metrics between the two DSS groups. This was found to be the case in Experiment I (Veitch et al., 2018) and Experiment II (Thistle et al., 2019) and was expected to hold true here. In Experiment I and Experiment II, differences in performance were found between the results of the low and high ice concentration groups, but these differences did not interact with the factors of experience and training. For example, in Experiment I, the Seafarers group performed significantly better on average than the No DSS group in both mild and severe ice concentration conditions. This can be explained by their higher level of experience, leading to better results regardless of ice conditions. It is expected that the DSS group will also perform better than the No DSS group in both ice condition levels.

2 Literature Review

A literature review was performed using Google Scholar's general search tool, and Memorial University of Newfoundland's library OneSearch tool with the key search terms including: *Decision support system; decision support system marine navigation; offshore; maritime navigation; maritime environment; full factorial experiments; ice navigation*; and *ice management.* A preference was made to include relatively recent resources, published within the last 10 years. However, exceptions were made for highly relevant, interesting, or unique works. This literature review aimed to provide an overview of the current status of research on ice management, decision support systems, and statistical analysis of experiments. A summary of the key literature is presented in this section.

2.1 Offshore Ice Management

2.1.1 The IMO Polar Code

Maritime navigation in polar waters has been regulated by the International Maritime Organization (IMO) *International Code for Ships Operating in Polar Waters* (the Polar Code) since its ratification on January 1st, 2017 (International Maritime Organization, 2010). This document defines the limits of Polar Regions, and outlines the requirements for the construction and operation of a vessel so that it can operate safely within the bounds of its given ice class. This includes hull design, control systems, emergency egress, lifesaving equipment, personnel competence, and stability, amongst others. The IMO in conjunction with Transport Canada has also developed a model course for basic training of ship personnel operating in polar waters within the requirements of the Polar Code. There is a requirement for vessels transiting in defined polar regions to have a certified ice navigator on board, a position that can be acquired by a deck officer through the completion of an approved training course (International Maritime Organization, 2017). This training provides an overview of ice characteristics, regulations, and vessel characteristics for manoeuvring in ice, passage planning, ice breaker assistance, crew preparations for safety, and the environment. This is intended to help standardize and improve the safety of Arctic navigation (International Maritime

Organization, 2017). It does not, however, support or train for industry-specific ice management operations such as those that are done in the offshore oil industry.

2.1.2 Ice Management

The focus of this research is on simulated emergency ice management for offshore oil installations on the Grand Banks of Newfoundland. Ice management is defined in the International Organization for Standardization (ISO) standard ISO 35104 as:

"*…the sum of all activities, carried out with the objective to mitigate hazardous situations by reducing or avoiding actions from any kind of ice (sea ice or glacial ice)… "* (International Organization for Standardization, 2018).

Although the Grand Banks region is well south of the Polar Regions as defined in the Polar Code, this region nonetheless has frequent incursions of pack ice that travel south from Polar Regions, and the Labrador Sea (Dunderdale & Wright, 2005). Pack ice has the potential to push floating installations off-station, subsequently causing operators to halt oil production to mitigate risks and environmental damage. An incursion of pack ice can also negatively affect the ability of lifeboat evacuations in the event of an emergency (Simões Ré & Veitch, 2007) As such, offshore supply and standby vessels may be tasked to clear ice from the area (Dunderdale & Wright, 2005).

The Canadian Coast Guard is an organization that regularly performs ice management for commercial operators, and in some cases, for emergency response. These activities range from ship escorts through ice-covered areas, to breaking ice in rivers. The Canadian Coast Guard College has recently begun developing and implementing lessons for specific ice management activities in a simulator (Canadian Coast Guard College, 2018). This allows future and current navigation personnel to practice ice management activities with less risk than would be associated with performing them in the field.

2.1.3 Decision Support Systems and Automation

The framework for the DSS used in this experiment drew heavily on the work of Smith et al. (2020). In their research, a DSS was created to transfer ice management knowledge of expert seafarers and to present it to new operators (Yazdanpanah, 2021). This expert knowledge was captured through interviews and simulator sessions with experts who critiqued the video performance of seafaring cadets and performed ice management scenarios of their own in a simulator. A DSS was then created using a modified case-based reasoning algorithm. A knowledge base of ice management strategies was used to classify the approaches of new users when assistance was requested. These strategies were then matched against previous cases in a case base, and the closest match was presented to the user with instructions on how to implement the ice management technique (Yazdanpanah, 2021, Smith, et al., 2020). This research was critical to the development of the DSS that was adapted for use in this current research.

The use of an onboard DSS specifically for ice navigation is thought to be novel. Previous studies have investigated the effectiveness of a DSS for marine navigation. For example, Nilsson et al. (2008) performed a simulation study of an advanced DSS for ship navigation. In this experiment, 39 Master Mariner participants with experience levels ranging from < 1 year, to > 30 years were recruited. The participants were tasked to navigate a simulated vessel through a busy Hong Kong fairway. The participants were split into two groups: an advanced bridge group featuring an Electronic Chart Display Information System (ECDIS) as well as a conventional chart, and a conventional bridge, with only paper charts available. Additionally, the advanced bridge allowed for radar overlay, and displayed curve heading indicators which automatically compensate for the vessel's drift due to wind and currents. Hong Kong was chosen as the location since it was thought that it would be less likely that the participants were familiar with this region than one from their local area (Sweden). All participants were provided with habituation and were then randomly assigned to a bridge (conventional or advanced). Overall, no significant differences were observed, however there was a tendency for the experienced navigation officers to perform worse using the advanced bridge. Furthermore, the conventional bridge was rated to have a higher workload than the advanced bridge by participants (Nilsson, Gärling, & Lützhöft, 2009). To this point, other studies based on real world accidents have pointed out the negative impact the over-digitization of a marine bridge can have on safety (Mallam et al., 2014).

A study by Pietrzykowski et al. (2017) summarized the research to date on a Collision Regulation (COLREG) based navigational DSS created by a team at the Szczecin Maritime University. This Navigational Decision Support System was shortened to the acronym NAVDEC. This system has been validated in both simulator and ship-based environments but has faced challenges in its development and implementation. The study discusses the process behind navigational decision making and outlines the large number of information systems currently available to operators, including Electronic Chart and Display Information Systems (ECDIS) and Vessel Traffic Services (VTS). These systems do not provide decision support guidance to operators, that is, they provide the operator with a large quantity of data with which to make a decision themselves. Advances in artificial intelligence (AI) have given rise to the ability for route planning and optimization software to be developed with sufficient power to solve multifaceted problems. The NAVDEC software integrates these information systems to provide collision avoidance manoeuvers to the responsible navigation officer on the bridge. This helps to prevent accidents caused by human error by recommending a course and speed which will avoid a target and will pass at a pre-specified distance in a fashion that complies with the COLREGS. This does not relieve navigation personnel of their responsibilities, but it helps to guide a decision. The NAVDEC software presents itself in a simple graphical user interface (GUI) with a mixture of numerical, pictorial, and map based guidance to suggest an optimal course for collision avoidance. Simulation studies demonstrated generally positive results through the accurate calculation of closest point of approach (CPA) and time to closest point of approach (TCPA) (Pietrzykowski, et al., 2017). Additionally, the system was tested on a cargo vessel sailing from Spain to several West African ports. This test found that 70% of calculations for CPA and TCPA performed by the NAVDEC system matched the calculations performed by the traditional Automatic Radar Plotting Aid (ARPA) system. Although no specific study was done to compare collision avoidance performance of seafarers using NAVDAC against seafarers without, it remains a promising technology that will likely continue to be developed. The task of collision avoidance differs from the more abstract and variable ice management tasks which have different performance metrics and goals, however commonalities remain in the GUI design and ultimate goal: to reduce human error at sea (Pietrzykowski, Piotr, & Borkowski, 2017).

Another study from 2019 examined the need for maritime training to include a greater focus on how automation uses data, and what the data are used for (Westin, et al., 2019). This research examines the ways in which advanced automation has contributed to several marine accidents involving modern bridge systems. Using a field study and literature review of maritime accidents, it was found that difficulties in understanding the maritime path predictor contributed to incidents when the predictor failed due to errors in position data. Increasing the transparency behind how the data is being used and its limitations could be valuable. This finding was used when developing the DSS briefing so that participants understood the limitations behind the technology.

A 2018 study examined the limitations of automations and industry trends to improve human-automation partnership (Pazouki, et al., 2018). This study reviewed recent aviation accidents to highlight the influence of automation on human performance, and attempted to apply this to a maritime context. In this study, an experiment was performed with 12 deck office cadets, showing them a video of a simulated scenario. The scenario shown initially was run in a nominal way to build trust in the participants. Following this, subtle faults of the automated systems were introduced to gauge whether the participants recognized the failure. The faults chosen were “possible traffic conflict” alarms, and “deviation from course” alarms. These alarms would function properly initially, but would fail in later scenarios. The participants were instructed to make notes of when they thought the systems were not functioning as intended. Half the participants were provided with training in the form of a passage plan briefing with paper charts, and the half were not. This was to test for the influence of training on the response. The participants were also asked to subjectively mark their trust in automated systems. A tendency was found that most participants had a high level of trust in automation, however there was no strong correlation to the time it took to recognize failure. There was also no strong correlation with whether the participants were from the Trained group or not (Pazouki, Forbes, Norman, & Woodward, 2018).

Tsou et al. (2010) took a different approach, incorporating a genetic algorithm for ship collision avoidance route planning. This algorithm computed a path based on a CPA of two vessels and calculated the shortest route to avoid a collision risk. In this research, a virtual agent navigated a ship and successfully avoided a collision. Testing involving humans was not performed (Tsou, et al., 2010).

Skjong et al. (2019) published a conference paper studying a modular DSS that was relatively customized for multiple marine applications. For the purposes of the research, they examined data driven, and model driven categories of DSS for marine applications. Four marine based cases were studied: offshore crane operations, payload vessel interaction, and two different types of trawling operations. The data driven DSS described by Skjong et al. (2019) is a similar architecture to the DSS used in the current research in Experiment III, although they do not use a case-based reasoning algorithm. The authors of that study found that the data driven model coupled well with a GUI to inform decision makers to how to optimize their operations (Skjong, et al., 2019).

The visual representation of data in a display has a high degree of importance from a human factors perspective. Chapter 8 of the book *Designing for People: an Introduction to Human Factors Engineering* (Lee et al., 2017) discusses some of the ways a navigational display can be designed such that it is optimal for human use. Although human factors design was not formally considered in the development of the DSS in this experiment, the comments of the participants in their exit interviews could be analyzed from a human factors standpoint to see what worked well, and what could be improved on. Comments from participants on desire to have tactile control, different colours highlighting important components, and a less cluttered display were cross-referenced with *Designing for People* to improve understanding as to why the participants felt this way, and how this could be improved in future.

2.1.4 Design of Experiment and Analysis

The design of the current experiment was guided by the works of two previous graduate students, who performed similar experiments using the ice management simulator at Memorial University. In the study by Veitch et al. (2018), a 2^2 full-factorial experiment was performed using inexperienced cadets, and experienced seafarers to measure the effects of seafarer experience and ice concentration on ice management performance. Veitch et al. (2018) used a formal experimental design using two levels for each group. The experimental design made use of Montgomery's *Design and Analysis of Experiments* in the planning process, which was repeated in the current study as well (Montgomery, 2013). The low and high levels for the ice concentration factor included mild ($4/10^{th}$) and severe ($7/10^{th}$) ice conditions. The low and high levels for the experience factor included inexperienced cadets and experienced seafarers. Veitch et al. (2018) found statistically significant evidence that experience had an effect on ice management performance, and that there was over 80% design power supporting this finding. According to Montgomery (2013), this is generally considered to be an acceptable design power, which means that there is an acceptable sample size to accept the experimental findings. Veitch et al. (2018) also found that there was less variance in the results of the experienced seafarers and that some of the cadets performed well – better even than the mean seafarer response. Ultimately, this study provided a method with which to perform an effective study on factors influencing ice management. Further, the experiment by Veitch et al. (2018) provided a control group of inexperienced operators who had no navigational aids, such as the DSS used in the current work. Additionally, the experienced group of seafarers in the study by Veitch et al. (2018) provided an upper bound on performance that could also be used in comparisons with the DSS group.

Thistle and Veitch (2019) built on the work by Veitch et al. (2018) by testing the effects of training on ice management performance. The goal of this research was to both analyze the effects of training, and to develop a method to predict the amount of training required to achieve the desired level of ice management performance. This work saw two groups of inexperienced cadets provided with one or two sessions of ice management specific training. The training program was produced using the work of Veitch et al. (2018), and the IMO Polar Code (International Maritime Organization, 2017) to train cadets using the best practices of ice management for the task. The study found that the cadets' performance subsequent to training improved in all metrics, and that in some metrics, a second training session significantly improved the results. When developing the DSS, portions of the training session developed by Thistle and Veitch (2019) were incorporated into the DSS instructions assist DSS users. Additionally, Thistle &Veitch (2019) made use of a similar experimental design to that of Veitch et al. (2018), with two levels of training, and two factors (training and ice concentration). When comparing results between different experimental campaigns, the assumption of randomness is violated, necessitating that the experiment be considered a quasi-experiment (Shadish, Cook, & Campbell, 2002). Thistle &Veitch (2019) acknowledged this when comparing their results to the results of Veitch et al. (2018), and the same limitation is also applied to the current experiment. Results from experiment II were analyzed using methods for nonparametric statistics, which are defined as data that are not based on standard statistical distributions, for example, the Normal Distribution. The results of the current experiment risk violating several rules required to be analyzed as a parametric equation, and thus analysis as a nonparametric statistic must be considered (Corder & Foreman, 2014).

Cohen (1992) discusses various statistical methods of calculating the design power, a method of testing the statistical strength of conclusions. Cohen discusses a value of *d*, which can be used for the power calculations. This *d* value is also known as the effect size, which is a function of the difference in means divided by the common standard deviation of two groups. The effect size can also be thought of as the discrepancy between the null hypothesis, and the alternative hypothesis, according to Cohen (1992). Cohen discusses how having insufficient design power, generally thought to be less than 80%, can result in what's known as a Type II error where the null hypothesis is incorrectly maintained (Cohen, 1992).

3 Methodology

In this section, the experimental design and procedure are discussed in detail. First, the process of the formal experimental design is described. The participant selection process is then described. Following this, the simulator is described in detail, the task given to the participants is explained, and the experimental procedure is summarized. The design and operation of the DSS are then explained and finally, the analysis procedures are described.

3.1 Design of Experiment

The experimental design chosen is a 2^k full factorial design using nine replicates, and k = 2 factors (Montgomery, 2013). The experiment designed in a similar way to that of Experiment I and Experiment II to allow for datasets from both to be used in these comparisons, and as the control group. In Experiment I, the two factors studied were experience and ice concentration (Veitch et al., 2018). Experiment II examined the effects of two levels of training on inexperienced cadets, and ice concentration (Thistle & Veitch, 2019). In the current research, Experiment III, the effects of DSS guidance and ice concentration were examined. As the experimental procedure was closely modelled after Experiment I by Veitch et al. (2019), the data from the cohort of cadets without access to a DSS could be used as the direct control group (the No DSS Cadets). Additionally, Experiment II by Thistle et al. (2020) was designed in a similar fashion, so the results from that experiment, which examined the effects of training, can also be compared against the effect of the DSS. In the current experiment, studying the DSS as a factor enables comparisons against the factors of Experiments I and II.

As experiments I, II, and III were completed at different times and with different researchers, the randomization of the analysis and assumption of independence are not valid, thus, it was considered a quasi-experiment (Shadish, Cook, & Campbell, 2002). Steps were taken to minimize differences between the experimental phases, including the use of scripts so that briefings to participants were consistent, use of the same habituation scenarios, and the same experimental equipment including the simulator as in the first two experiments. Additionally, Experiment III was treated as a split-plot design, a method that can be used to account for differences of participant groups between experiments (Montgomery, 2013), and to account for the lack of randomness. Logistical challenges also necessitated a departure from the standard protocols of a Completely Random Design (CRD) for the cohort studied in Experiment III.

Design power of a statistical test is defined by Cohen as the probability of achieving statistically significant results (Cohen, 1992), that is, the probability that the null hypothesis, H_o will be rejected when it is false. Design power is calculated using three variables. These are the significance level, α, the sample size, n, and the effect size. A desired design power can be used to calculate the required sample size, N, of an experiment to achieve a sufficient probability of statistically significant results based on a hypothesized effect size. The effect size index for different independent means in a *t*-test is referred to as *d* (Cohen, 1992). It is normalized by taking the difference in mean populations, and dividing them by the standard deviation. Cohen's *d* value can be used when the data are normally distributed. Cohen's *d* values between 0.2 and 0.5 are considered to be small, a value from 0.5-0.8 medium, and a value above 0.8 is considered to be a large effect.

Experiments I and II estimated that nine replicates, or 18 participants per group would be required to achieve a design power of 80% when $\alpha = 0.05$ (Thistle & Veitch, 2019). A post-hoc power analysis confirmed that this was appropriate and provided at least the 80% power desired to correctly reject the null hypothesis without significant concerns of a type II error. Based on these results and in the absence of additional data for the DSS efficacy, a hypothesis was made that in Experiment III, the DSS group performance would be significantly better than that of the No DSS group, approaching the results of the Seafarer and Training groups. The result of this would be an effect size of the DSS group when compared to the No DSS group that roughly equals that of the Seafarer and Training groups over the No DSS group. As such, the same power requirements were estimated as in Experiments I and II, so the Experiment III was designed to include 18 participants as well.

3.2 Participants

Twelve nautical science seafarer cadets, and 6 post-secondary students in other programs were recruited. All were considered to be similarly inexperienced for the purposes of this experiment. A summary of their academic and professional marine experience can be found in Table 24 and Table 25 in the Appendix. Due to a saturation of the nautical science recruitment pool, only 12 participants could be recruited. This number of participants was not found to provide sufficient design power in a post-hoc analysis. To improve the design power, it was decided that the inexperienced seafarer category that had previously been limited to nautical science cadets should be expanded to allow for other students to be recruited outside of this program to reach 18 participants so that consistency with Experiment I and Experiment II could be maintained, and design power could be improved. It was assumed that nautical science students, many of whom were recruited in their first several weeks of their program, were not likely to differ in knowledge or experience significantly from a student from the wider academic community. To reduce any discrepancies, students from the general academic community who had some marine related experience were recruited for the second pool. These will be referred to here as "irregular" participants. In total, six "irregular" participants were recruited from the wider student university student population. One of these had experience on pleasure craft, and one had three weeks of commercial marine experience. This selection process ensured that the six "irregular" participants were more likely to be aware of some basic maritime-related terms, such as bow, stern, port, and starboard.

3.3 Simulator

The experiment was performed using an ice management simulator located in the Faculty of Engineering and Applied Science at Memorial University of Newfoundland and Labrador. The simulator was built for marine safety research, specifically for scenarios involving sea ice. The simulator is immersive, featuring a 360° screen that displays images of 11 projectors. The diameter of the screen is approximately 8m. In the center, a simplified bridge console provides basic controls for a simulated anchor handling offshore supply vessel (AHTS). The vessel is modelled after a vessel of 75m in length typical of Newfoundland and Labrador's offshore oil industry on the Grand Banks. The simulated vessel has twin 5369 kW diesel engines coupled to fixed pitch propellers, and is provided with 895 kW tunnel thrusters at the bow and stern. A schematic of the simulator can be seen in Figure 2.

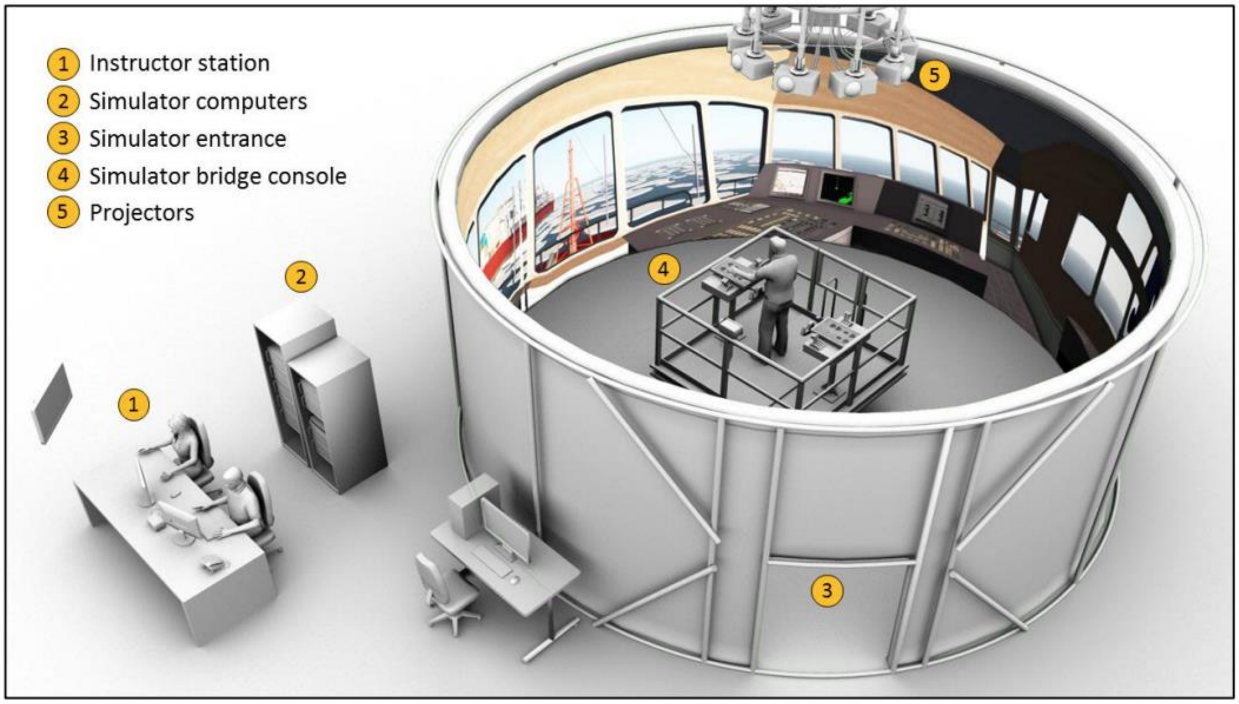


Figure 2: Schematic of simulator showing the various components. 1) shows the instructor station. 2) shows the computers which run the simulator's physics engine and projectors. 3) shows the entrance to the simulator. 4) shows the bridge console of the simulator. 5) shows the simulator projectors. (Veitch, et al., 2018)

The bridge console consists of a small platform with identical fore and aft-facing controls. This allows for the participant to choose whether they want to handle the vessel in a forward or aft-facing position, a practice that is common for operations with AHTS vessels. A view of the controls as seen by the participant can be seen in Figure 3. These controls are simplified, and consist of two throttles for port and starboard main engine control, a bow and stern thruster control, and a ship's wheel for rudder control. A display is used to present information to the participant regarding vessel speed, engine speed, heading, and rudder angle. No radar or chart plotter is provided. The participant can use a VHF radio to request distances and bearings off objects from the experimenters. The experimenters sit outside the simulator theatre at the instructor station. At the instructor station, the experimenters have access to this information on their display, which they can relay to the participant as if they were a crew member on the bow or bridge wing. The DSS is positioned to the left of the participant when in the forward-facing position, and is presented on a laptop (Figure 3). It can be activated by using the laptop's trackpad to press the on-screen 'assist' button.

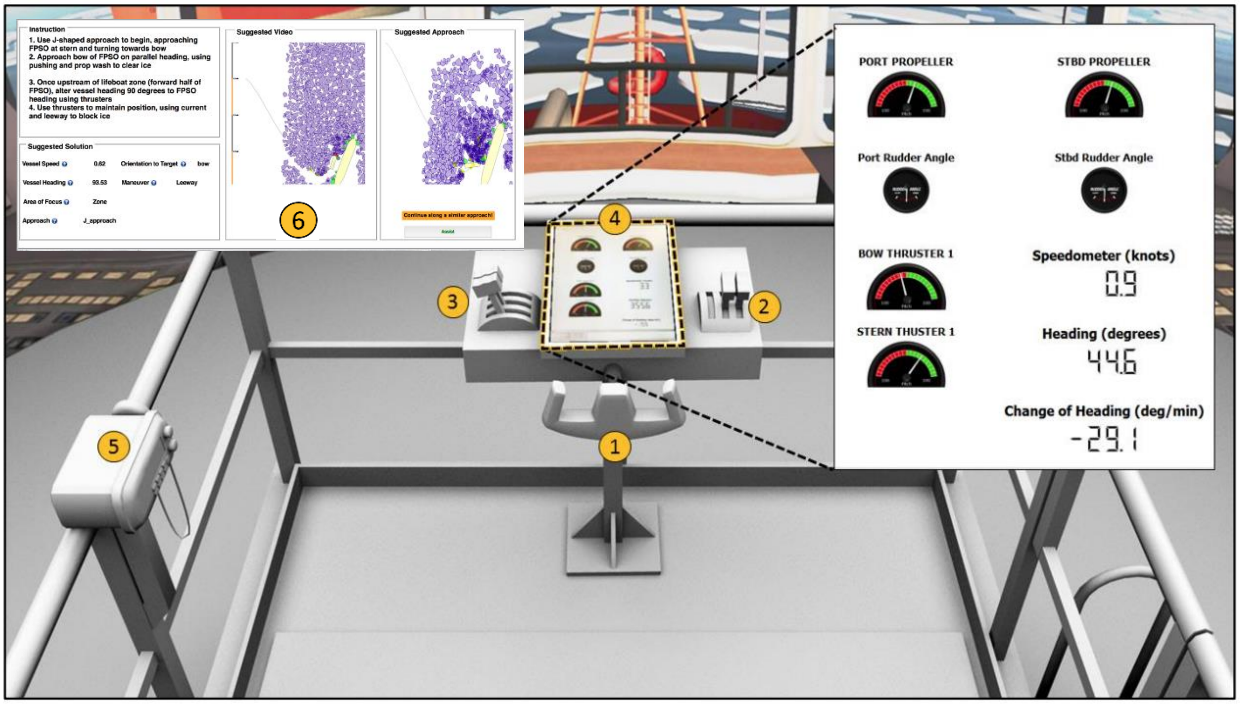


Figure 3: View of controls from Participant's viewpoint. 1) shows the ship's wheel. 2) shows the throttle controls for the propellers. 3) shows the throttle controls for the side thrusters. 4) shows the primary display. 5) shows the marine radio for communication to the instructor station. 6) shows the DSS display, approximately in the same position as in the simulator, although it is not to scale here (Veitch, et al., 2018)

3.4 Task

The scenario for the experiment consists of an ice management operation for a floating production, storage, and offloading (FPSO) vessel in an ice field with a 0.5 kt current. A schematic of this scenario can be seen in Figure 4. The ice is considered to be medium first-year ice with a thickness of 0.4m - 0.7m and a concentration of either $4/10^{th}$ or $7/10^{th}$, depending on the randomized assignment of the participant. Concentration is defined as the ratio of ice cover to open water and is expressed in units of tenths ($X/10^{th}$). The task that was given to the participant was to use the supply vessel they are piloting to clear ice from the lifeboat launch zone of the FPSO. Various metrics are used to quantify their performance, but ultimately, participants are instructed to clear as much ice as possible for as much of the scenario as possible from the FPSO's lifeboat launch zone. This scenario is modeled after a plausible ice management assignment for a standby vessel working in the Grand Banks region.

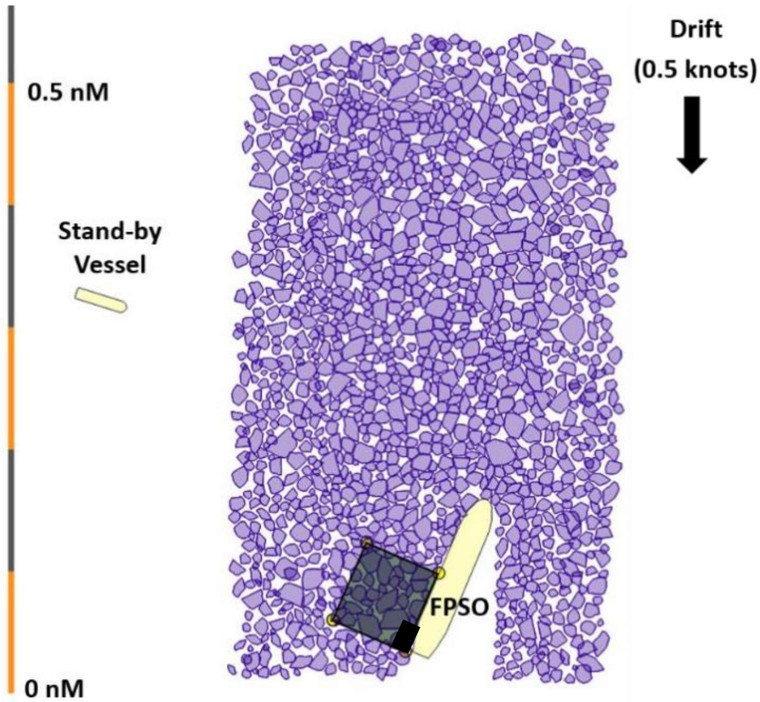


Figure 4. Schematic of Emergency Ice Management scenario used in the experiment, showing the zones to be cleared. The figure also shows the participant's ship, the FPSO, the direction and speed of the current, and the ice floes present.

3.5 Experimental Procedure

This experiment was reviewed and approved by Memorial University's Interdisciplinary Committee on Ethics in Human Research (ICEHR), with file number 20220482-EN. The approval letters and documents submitted to the ethics board can be seen in the appendix in 7.3. As discussed in section 3.1, the experimental procedure was designed after the protocols in Experiment I and Experiment II so that quasi-experimental results could be compared between all groups. The simulator, scenario, habituation script, and participant qualification criteria used in all three experiments were the same. A detailed description of the experimental procedures for Experiments I and II can be found in the literature by Veitch et al. (2018) and Thistle and Veitch (2019).

For recruitment, an email was circulated to the students in the local nautical science program, and posters were placed on the walls. Additionally, a visit was made by the researchers to the classrooms of all years of students to explain the experimental procedure and advertise recruitment. Participants were provided the contact information for the researchers and were invited to contact if interested. Participants who contacted the researchers were scheduled into a time-slot that was compatible with the schedules of both parties. They were then randomly assigned to an ice concentration group on the day of their simulator session, but were not informed which ice concentration group they would be placed in before arriving.

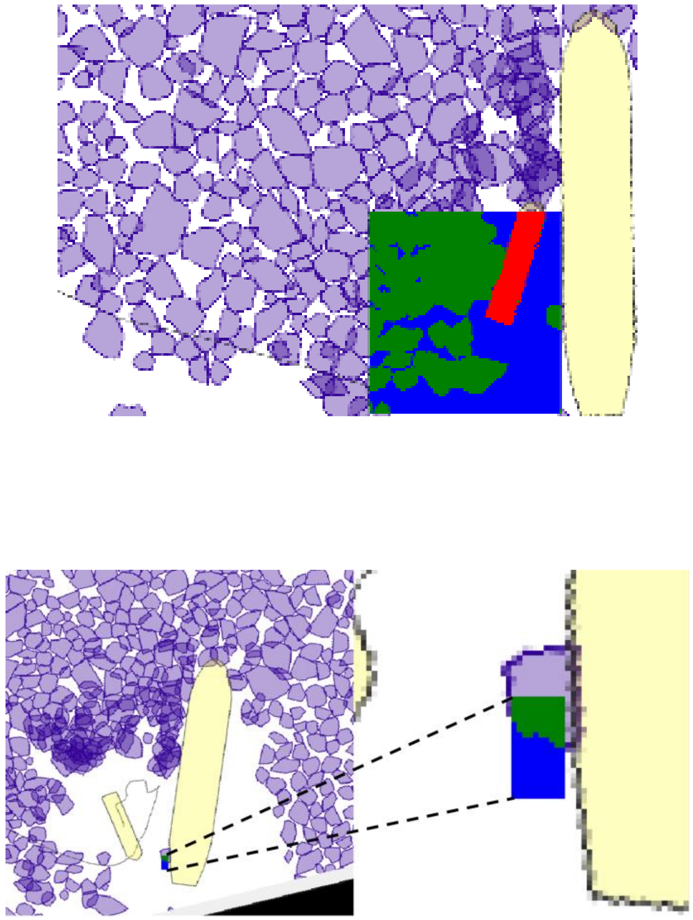
Participants were then provided with an informed consent form, and were asked to complete a questionnaire to determine their level of experience. This captured information about their time spent in formal maritime education, time at sea, and time spent working in ice. To ensure safety, the participants were asked to complete a simulator sickness questionnaire to establish a baseline of their current physical state before entering the simulator. Some common symptoms they were asked to be aware of are nausea, headaches, or dizziness. This was repeated periodically to ensure participant safety, and participants were informed that they could stop the experiment at any time should they no longer feel comfortable. No adverse effects were reported by any participants.

Upon completing the questionnaire, participants were shown the simulator controls and were given three different vessel habituation scenarios to complete. This was intended to familiarize them with the controls in the simulator and the virtual environment, but was not intended to provide them with ice management training. These habituation scenarios were identical to those of Experiments I and II to ensure all participants across all cohorts had the same level of familiarity with the facility prior to beginning the experiment.

Following the habituation, participants were given the task required by the experimental scenario, and were shown how to use the DSS. The DSS habituation gave them the opportunity to be aware of the assist function, gain knowledge about the information presented to them, and develop an understanding of the strengths and limitations of the technology. The latter point is important to ensure that the participants do not misunderstand the information being presented to them, which could result in incorrect use of the technology (Nilsson, et al., 2009).

The experimental scenario was kept identical to Experiment I and Experiment II to ensure consistency for comparisons across groups. The purpose of the experiment was to use the AHTS under the participant's control to perform an ice management operation where ice was to be cleared from two zones on the FPSO's port side. Figure 4 provides an example of the 7/10th ice concentration scenario. The two zones to be cleared can be seen, with the larger ice management zone shown as a semi-transparent square on the port side of the FPSO, and the lifeboat launch zone shown as a smaller black rectangle inside this. This shows the initial position of the participant's AHTS stand-by vessel, in the top left. There is a 0.5 kt current pushing the ice south.

The larger zone is a 120m x 120m square area. It leads from the stern of the FPSO to approximately amidships (Figure 5a). The lifeboat launch zone is 16m x 8.2m, and is directly underneath the FPSO's lifeboat (Figure 5b). The instruction given to participants was to clear and maintain the ice free status of both these zones to the best of their abilities over the course of the 30 minute scenario. The scenario is derived from realistic ice management activities used in regions where the presence of sea ice is likely. The Newfoundland Grand Banks are an example of such a region.



a)

b)

Figure 5. Image processing of ice concentration by pixel count. a), above, shows the larger ice management zone with the AHTS inside it, while b), below, shows the lifeboat launch zone.

Participants were informed that their score would be based on reducing the quantity of ice as much as possible in the larger ice management zone, and the amount of time that the lifeboat launch zone is completely clear of ice.

The simulated ice floes were generated by randomly sampling a lognormal distribution of ice. The floes were given a uniform thickness of 40cm (Veitch et al., 2018). The FPSO was static throughout the scenario, and did not yaw, pitch, or roll. The ice drifted at 0.5kt with the current from the port bow of the FPSO and flowed past the stern. Participants were informed that they had the option to use the DSS as many or as few times as they wished. In the DSS habituation, participants were informed that the software evaluated their current position, heading, and orientation to the FPSO and used this to provide them with strategic recommendations for a tactic or approach to follow.

The participants were informed they could use the marine radio to request distances and bearings of targets, such as ice floes and the FPSO, from the simulator operator, who had access to distance and bearing data off objects. They were told to envision that they were the master, and the simulator operator was a crew member on the bow, stern, or bridge wing of the vessel who could provide distance estimates. Other than this, the participants were not provided with any additional guidance. After 30 minutes, the experiment ended.

Upon completion of the scenario, a final simulator sickness questionnaire was provided to the participants. They were then given an exit interview to allow for qualitative analysis of their impressions of the DSS efficacy. The participants were asked questions about the reasoning behind their chosen strategy, their perceived score on a scale of 1-5, and the usefulness of the DSS on a scale of 1-5. They were also asked which specific components of the DSS they found helpful, which they did not use, and which they thought could be improved. Before departure, participants were asked to refrain sharing the contents of the experiment with their peers to avoid influencing future results should one of those peers chose to participate.

3.6 Decision Support System

The DSS used in this experiment was developed by a previous graduate student at Memorial University (Yazdanpanah, 2021). It is an example of first-degree automation as defined by the IMO, meaning seafarers are provided with on-board decision support, but are still fully responsible for decision making (International Maritime Organization, 2021). This is considered a low level of automation. The DSS uses a case-based reasoning (CBR) model, which consists of three different components. These are the case base, knowledge base, and the graphic user interface (GUI) the last of which presents the user with information and allows them to interact with the system (Kolodner, 1992). Yazdanpanah (2021) used expert knowledge gathered in a study by Smith et al. (2020) to populate the knowledge base and to classify the case base. The *case base* is a database made up of previous experimental runs from Experiment I and Experiment II. Next, the *knowledge base* was created by interviews with experienced mariners in a knowledge capture study. The experienced mariners were asked to critique the performance of participants in these runs, and to identify the most critical factors for effective ice management (Smith et al., 2020). The factors identified included vessel particulars (heading, speed, specific ice class), ice conditions (ice concentration, floe size, ice type and thickness), and task objective (area of ice to be cleared). These factors were ranked by importance by the expert seafarers for classification of cases in the case base. They were also asked to participate in a simulator session so that their recommended approaches could be tested in the same setting. The knowledge gained in the interviews was also used to provide expert suggestions to improve performance for users of the DSS in Experiment III. Finally, the GUI was created to present this information to new users.

A method was created to classify specific approaches using factors such as position and heading in relation to the FPSO and match those to a previous case which should be emulated (Yazdanpanah, 2021). In general, having more cases in the case base improves the ability of the DSS to match specific cases and give good advice (Yazdanpanah, 2021). In practice, the case base contains approximately 40 cases that are considered acceptable. Cases that are considered below average were not included to avoid giving bad advice to participants. A DSS can be improved by adding new cases to the case base as more users participate, however this was not done for this experiment to avoid altering the properties of the DSS over the course of the experiment.

An example of the CBR algorithm can be seen in Figure 6. A typical CBR algorithm has four phases: the Case Retrieval, Case Reuse, Case Revision, and Case Retain. In this model, the Case Revision and Case Retain steps were excluded. When a new user requests assistance using the DSS, their current strategy can be matched to performances in the case base, and a case is retrieved. This is the Case Retrieval phase. The Case Reuse examines the reusable aspects and the differences between cases. Once matched, a case from the case base is shown to the participant as a case to emulate, with instructions and visuals on how to do so (Smith et al., 2020). This will provide them with guidance on how best to continue their approach to achieve strong ice management results. In a typical DSS, the Revise Case phase would alter the case being shown to the participant, however this was not done in this model. Also, typically, each case is retained in the Retain Case step so that the case base can be grown, however as discussed this was not done in this model to avoid changing the effects of the DSS between participants.

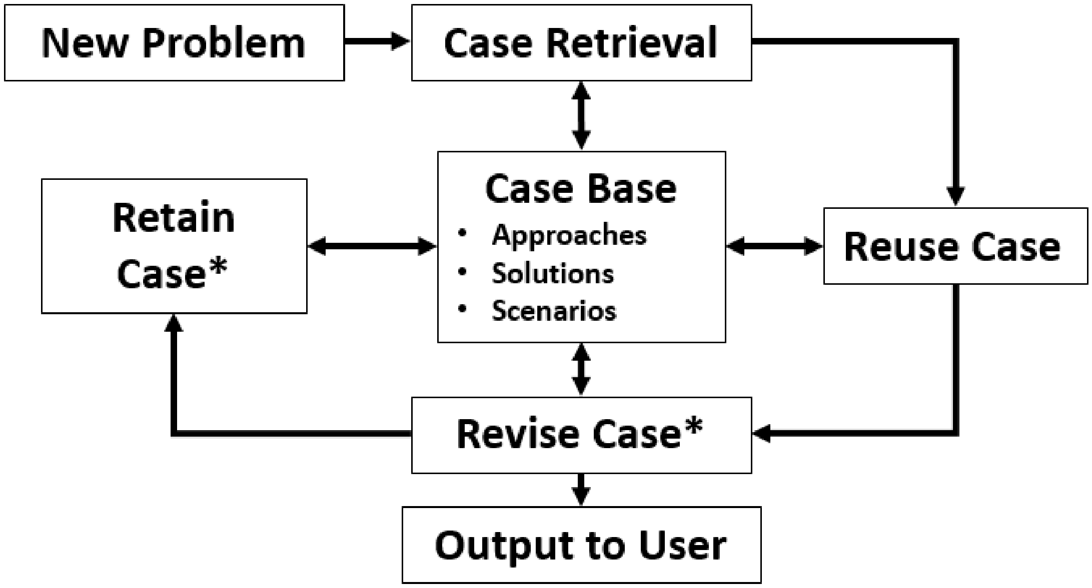


Figure 6: Case Based Reasoning model adapted from Smith et al. (2020). Features with a "*" are excluded from this model in practice

The final component of the DSS, the GUI, was altered from its original form as developed by Yazdanpanah (2021). Six individuals with various backgrounds were asked to complete an ice management task in the simulator using the DSS for pilot testing prior to beginning the actual experiment. They were then interviewed to see which components of the DSS were most useful, and which could be improved upon. Although a formal study on the effects and best practices of GUI design was not considered, this pilot test allowed for several incremental improvements to be made to the presentation of the assistance over the original design. It also allowed the CBR algorithm to be improved so that better decision support advice could be provided to participants. Since this pilot test was not formally a part of the experiment, no data were collected aside from qualitative notes for DSS improvements.

3.7 Modifying the Information Layout of the Decision Support System

An example of the initial form of the DSS GUI developed by Yazdanpanah (2021) can be seen in the Figure 7. The adapted version of the DSS GUI based on the user testing can be seen in Figure 8. The following section outlines the modifications made to the DSS GUI based on feedback from the pilot study.

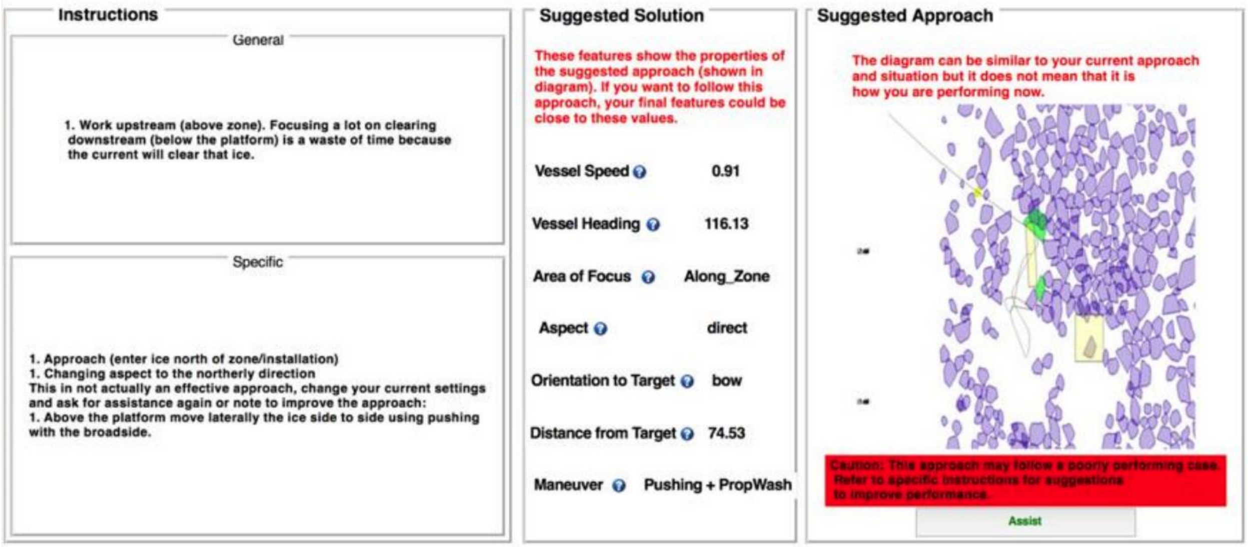


Figure 7: Original form of DSS GUI (Yazdanpanah, 2021)

Based on the comments from the pilot study, the information presented by the *Suggested Solution* in the center of the GUI shown in Figure 7 did not provide sufficient strategic advice. This information was more useful to a user once they were in the ice management zone. Further, it was thought that the large amount of information presented without much context could be confusing to an inexperienced user. This was confirmed by the first two participants in the pilot study. There was a perception that information presented in the center of the display is the most important, but the information contained in the *Suggested Solution* was not easily understood by new users, it led to confusion. This information could be useful to a more experienced user so it was not removed from the GUI. It was instead moved to the bottom left-hand side of the GUI, so that it was still available if desired, but was not presented as the first option for a participant to use. This can be seen in the bottom left of Figure 8. The *Suggested Solution* section provides specific operating parameters and strategy types. This includes a recommended heading, speed, orientation to the FPSO, and the name of the ice management approach suggested. The user can hover their mouse over the ice management approach for a detailed description, in case they are not familiar with the term.

The next component that was changed was the text-based instruction sections. Figure 7 shows that in the left-hand side of the display under *Instructions*, text is presented in both a *General* and *Specific* category. It was thought that these categories could be combined to avoid cluttering the display. Figure 8 shows that this change was completed, placing the text-based instructions on the upper left-hand side of the display. The instructions are presented in point-form, and were adapted from interviews with experienced seafarers (Smith et al., 2020). The instructions are customized for the specific approach being recommended.

There were several suggestions from participants in the pilot study that a live update of the vessel position overlaid onto the bird's-eye-view picture (seen on the right-hand side of the display in Figure 7 and Figure 8) would be helpful. However, this was not technically feasible with the resources available. In discussions with participants from the pilot study, it was thought that a video-replay could be a helpful way to deliver information, and would be clearer to a user than the *Suggested Approach* picture, or text alone. This was incorporated in the revised GUI and was placed in the middle of the display, as can be seen in Figure 8. A video of the suggested approach was presented to users in a way such that it replayed repeatedly in a loop. This was sped up 600 times so that the full 30 minute scenario could be replayed in six seconds. It was decided that six seconds was sufficient to avoid the user needing to take their attention away from their surroundings for too long and risk losing situational awareness, but was not so short that they were unable to process the information. The video allowed the participant to observe the dynamics of the ice flowing south in the current, and the interactions of the supply vessel with the ice. When combined with the written instructions, it becomes clear how the problem should be approached. This was thought to be the most important component based on observations from the pilot studies as it gave a clear overview of the recommended strategy. As such, it was placed in the center of the display so that it would be clear.

The user can request assistance as many times as required by pushing the assist button, shown in Figure 8 on the bottom of the right-hand panel. They might choose to do this if they're having difficulties with the strategy they're taking. The DSS will then match their current position, speed, and other parameters to the case base, and may suggest a new approach if these parameters differ from the last time assistance was requested. If there is not significant change, the advice will stay the same. For example, if the user requests assistance without changing their position, the advice will not be different.

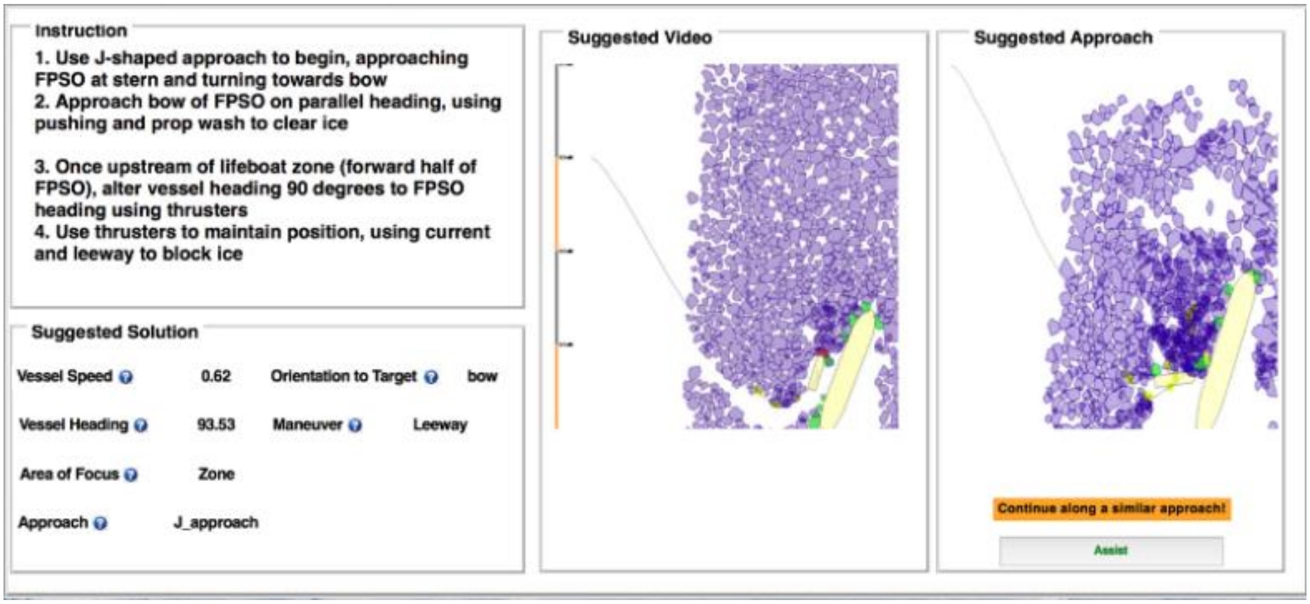


Figure 8: Adapted form of DSS GUI

When a user requests assistance within the first two minutes of a scenario, there is not sufficient data for the algorithm to work accurately. Originally, the solution to this was to present the user with a warning stating that there was insufficient data to give a good recommendation. The first two minutes is critically important in deciding the initial strategy, so it was felt that this was not acceptable in a DSS. This was altered so that within the first 30 seconds of the scenario initializing, when assistance was requested, one of three ideal cases representing three different approaches would be randomly shown. These strategies were either an approach from the bow of the FPSO, a direct approach to the lifeboat launch zone, or an approach from downstream of the FPSO's stern. An example of these approaches are shown in Figure 9.

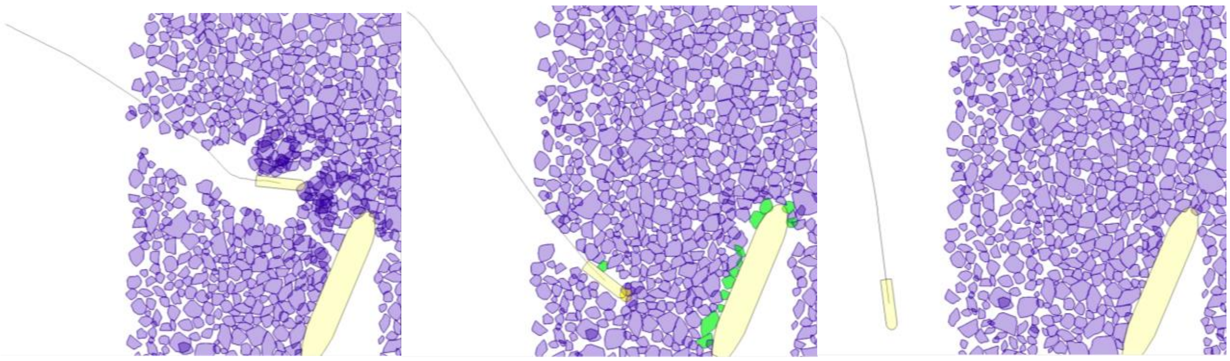


Figure 9: Example of three approaches: Bow, or Upstream approach (left), Direct approach (center), and Downstream, or J-approach (right)

After 30 seconds, but within the first two minutes, one of the three strategies initially shown randomly would be presented depending on the heading of the participant's vessel. For example, if the participant pointed their bow towards the stern of the FPSO and requested assistance, the DSS would recommend the downstream 'J-Approach', an ice management technique that has the responding vessel approach from the FPSO's stern, and turn towards the bow with a J-shaped course. This clears the ice from stern to bow. An example of this can be seen on the DSS display in Figure 8. This allows the participant to be presented with an ideal solution that will not be affected by the technical limitations of the DSS in the early stages. The ideal solutions were chosen due to their high performance scores, and by having clear strategies identified by experts during interviews by Smith et al. (2020). After two minutes, subsequent assistance requests from the DSS would utilize the CBR algorithm, with more variables in play for recommendations.

3.8 Analysis

To analyze the results, each participant's simulated performance was replayed in real-time, and screen captures of a birds'-eye-view from the instructor station were taken every second for analysis. An example of this bird's eye view can be seen in Figure 10**.**

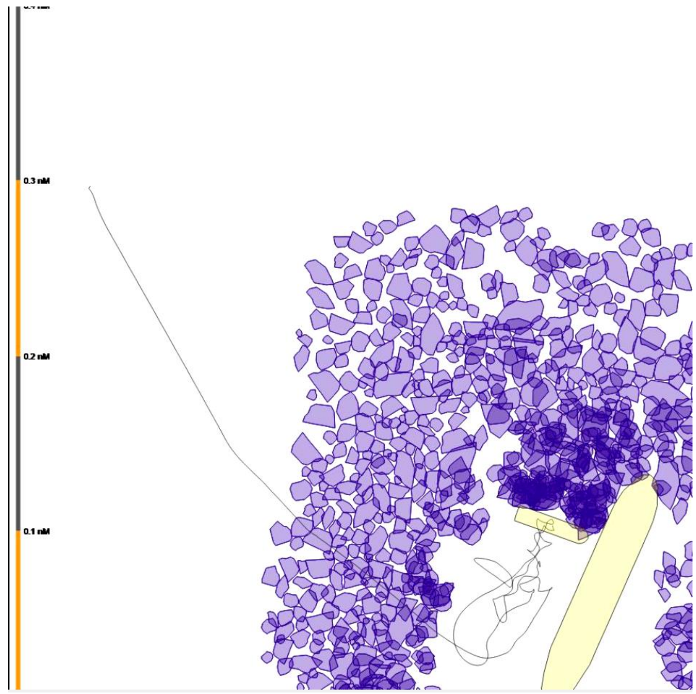


Figure 10: Bird's eye view of the run, as captured by screenshots

As the scenario played through, 1800 images were captured. Each run also generated a text file that was comprised of speed and position data for the participant's vessel. Image processing scripts were then used to calculate the specific ice concentration in the zones at regular intervals. An example of the image processing output can be seen in Figure 11. Pixels are classified based on whether they are the vessel, ice, or open water. This is used to calculate the concentration.

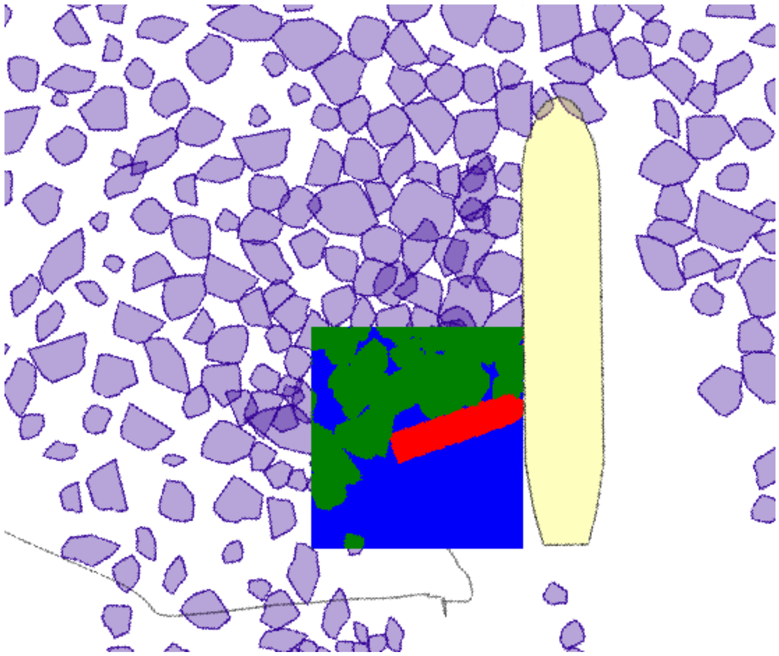


Figure 11: Sample of image capture script output. Open water is identified as blue pixels in the ice management zone. Ice is shown as green. The vessel is shown in red and is not counted in the ice concentration calculations.

As the ice was randomly generated when the scenario was first created, each participant started the scenario with the same initial condition. A baseline was generated by taking the average change in ice concentration over the zones if no ice clearing was performed. This baseline was compared against each run to generate data for each performance metric. Seven metrics were calculated to compare results between groups. Four of these metrics were taken from Experiment I and Experiment II, and three new metrics were added to capture more detailed data. The results from Experiment I and Experiment II were subsequently re-analyzed for these additional metrics. A summary of each metric can be seen in Table 1, and a more detailed explanation of each will follow. The *own-ship* refers to the simulated vessel controlled by the participant.

Table 1: Summary of performance metrics for participants

Metric	**Description**
Average Clearing	Average ice concentration reduction of ice in ice management zone
Maximum Clearing	Maximum ice concentration reduction in ice management zone
Total Clearing	Total ice cleared from ice management zone in units of area
Clearing to Distance Ratio	Ratio between average clearing and distance traveled by own-ship

Cumulative Lifeboat Launch Time	Cumulative time that lifeboat launch zone has a concentration of zero
Maximum Lifeboat Launch Time	Longest interval that lifeboat launch zone has a concentration of zero
First Time Clear	Earliest time in experimental run that lifeboat launch zone has a concentration of zero

The *average clearing* measures the mean change in ice concentration in the lifeboat launch zone. This metric was calculated by taking the difference in ice concentration in the larger zone shown in Figure 5a from the baseline at each 30s interval of the 30 minute scenario. Each interval was then averaged. Better performance was indicated by a larger value, which meant that more ice was cleared.

The *maximum clearing* metric measures the maximum reduction in ice concentration throughout the run. This would present itself as the lowest point in a plot of ice concentration and time. A larger value indicated a better result for this metric. One limitation of this metric is that a consistently low result is likely preferable to an instantaneously low result. A participant may score well in this metric as a maximum value but perform worse for the rest of the scenario. Depending on how short the duration minimum concentration is, it is possible that there would not be enough time to safely launch a lifeboat.

The *total clearing* metric is the summation of all incremental ice concentration drops, measured in units of km^2. This metric is cumulative, so a larger value indicates more ice was cleared in the scenario and is thus considered better.

The *cumulative lifeboat launch time* metric examines the smaller zone below the lifeboat, as shown in Figure 5b. When the zone was completely clear of ice with a concentration $0/10^{th}$ during one of the 30 second time steps, the lifeboat was considered able to launch. The total time in seconds that the lifeboat launch zone was completely clear of ice was then summed for the duration of the 30 minute run. A higher value was considered to be a better score since this meant that the lifeboat had the ability to be launched in open water for more time during the run. An example of the output for concentration in the lifeboat launch zone over the course of a single run is shown in Figure 5. The lifeboat can only launch when the value is at zero.

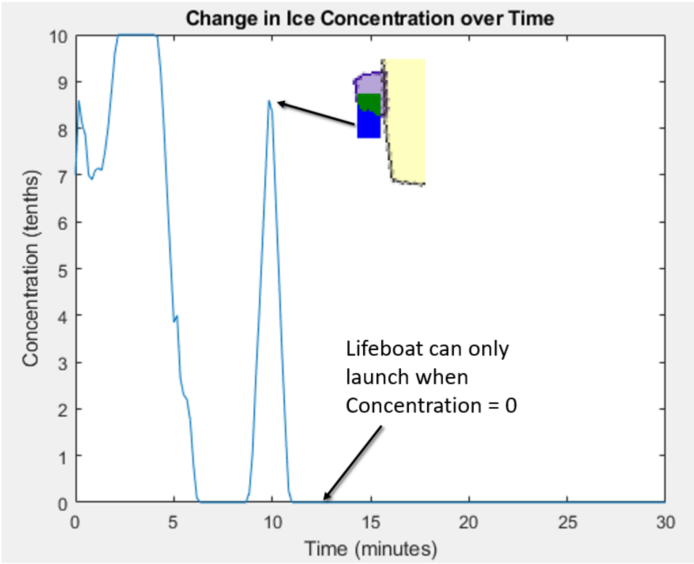


Figure 12. Example measurement of ice concentration in the lifeboat launch zone for a single run. A capture of the ice floe present in the lifeboat launch zone can be seen at the 10 minute mark to demonstrate how this affects the concentration metric. The lifeboat can launch when the concentration in this zone is zero.

The *maximum lifeboat launch time* metric examines the largest interval in which the lifeboat can launch. This is the longest stretch of time at which the concentration in the lifeboat launch zone is zero. Referring to Figure 12 for example, this would stretch from the 11 minute mark to the end of the 30-minute session for a total of 19 minutes, or 1140 seconds. A higher value is considered to be a better response, since in reality, it would allow a lifeboat crew more time to successfully launch.

The *first time clear* metric examines the earliest interval of time in which the lifeboat launch zone has a concentration of zero. Referring again to Figure 12, this would occur at approximately the six minute, or 360 second mark. In this case, a smaller value is considered to be a better response since it implies that a lifeboat will have the ability of launching sooner in the case of an emergency.

Finally, the *clearing to distance ratio* provided a look at the efficiency of a given participant's run. This was done by dividing the quantity of ice cleared with the distance traveled by the vessel. A higher value indicated more ice cleared per distance traveled and was considered to be a more efficient and controlled result.

A two-tailed *t*-test was performed when analyzing the results to ensure that all possible effects are checked for, whether positive or negative. Each result was compared to the DSS group and was checked for statistical significance, at 95% confidence.

Additionally, the effect size between the DSS and the control group was calculated so that design power calculations could be completed. The effect size is calculated the following way:

Equation 1: Calculation for Cohen's d. (Cohen, 1969)

$$d = \frac{\mu_{DSS} - \mu_{Control}}{\sqrt{\frac{\sigma_{DSS}^2 + \sigma_{Control}^2}{2}}}$$

where μ represents the mean value of each group, and σ represents the standard deviation. Cohen's d assumes that standard deviations are equal between groups. Since a perfect equal standard deviation is not likely to occur in the real world, a pooled standard deviation can be used, which combines the standard deviation from both groups. This is done by adding the square of each standard deviation, dividing them by 2, and taking the square root (Cohen, 1969). It is a unitless value.

4 Results

4.1 Effect Size and Design Power

In Table 2, the effect size for each comparison can be seen, using Cohen's *d* method (Cohen, 1969). This is averaged between the 4/10th and 7/10th groups. The results were all found to be approximately normally distributed when the residuals were plotted. This is explained further in section 4.2.1. This was then used in the calculation of the power for each comparison.

Table 2: Cohen's *d* effect size for both mild and severe ice concentration groups

Mean Effect (d)	Mean Conc Change	Max Conc Change	Mean Total Clearing	Max Time Clear	Cumulative Time Clear	Mean Clearing to Distance	First Time Clear
DSS to No DSS	0.67	0.19	0.68	0.16	0.25	0.84	0.28
DSS to Seafarer	0.89	0.44	0.90	1.14	0.95	0.94	0.69
DSS to T1	0.96	0.99	0.96	0.46	0.86	0.89	0.88
DSS to T2	1.42	1.01	1.41	1.25	0.78	1.38	0.64

Using the software G*Power, the post-hoc power was calculated for all metrics to see whether the effect size was sufficient to provide 80% power. An example of the power curve for the largest effect of the DSS when compared to the No DSS group is shown in Figure 13 below. There is an effect of 0.84 between these groups, with 18 participants in each group. From Figure *13* it can be seen that with a total sample size of 36, the power is 0.69, or 69%, which is below the desired 80% power.

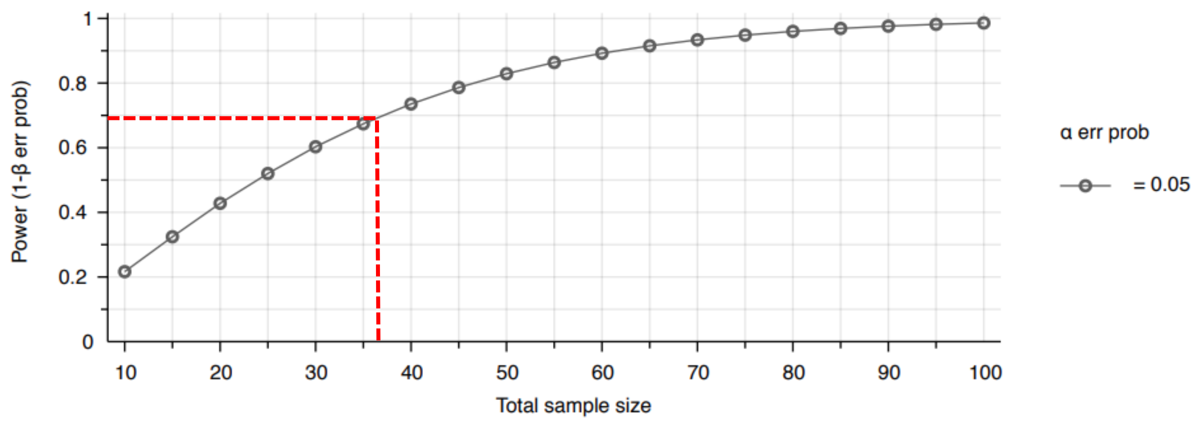


Figure 13: Post-hoc design power plot for clearing to distance metric of the DSS to No DSS groups

The design power for each metric can be seen in Table 3 below. From this, it can be seen that none of the results comparing the DSS to the No DSS group reach the desired power of 0.8, or 80%. This correlates to the effect sizes found in

Table *2*. From this, it can be seen that the smallest effects exist between the DSS and No DSS group. The highest power is found for the Clearing to Distance metric, at 0.69. This means that it is possible that a Type II error has been made, and that the null-hypothesis that no significant difference exists in between the DSS and No DSS groups has been falsely accepted. In metrics where the design power is greater than 80%, there is a high degree of confidence that there is a low risk of a Type II error.

In comparing the DSS group to the Seafarers, all metrics have a higher power, with five above 75%, and one above the desired 80%, at 91%. This is due to the higher effect size, due to the seafarer group performing better than the DSS and No DSS groups. Comparing the Training I group to the Training II groups, it can be seen that many of the metrics reach the desired 80% design power, with the Training II group having sufficient power in 5/7 metrics.

Table 3: Design power results

Power (from G*Power)	Mean Con Change ($X/10^{th}$)	Max Con Change ($X/10^{th}$)	Mean Total Clearing (km^2)	Max Time Clear (s)	Cumulative Time Clear (s)	Mean Clearing to Distance	First Time Clear (s)
DSS to No DSS	0.497	0.086	0.51	0.075	0.112	0.68	0.13
DSS to Seafarer	0.74	0.25	0.75	0.913	0.79	0.78	0.52
DSS to T1	0.8	0.82	0.8	0.27	0.71	0.74	0.73
DSS to T2	0.98	0.84	0.98	0.95	0.62	0.98	0.467

4.2 Primary Study Results Summary

The results of all 18 DSS participants when compared to all other groups are shown here. As discussed, six of the 18 DSS participants were recruited from outside the nautical science program. It was not expected that this would have a significant effect on ice management performance, and that it was preferable to have more participants to achieve higher statistical power, despite not being from the nautical science program. An analysis on the results when these different groups are compared will be presented in later sections.

Below, Table 4 and Table 5 compare results from all groups in experiments I, II, and III. Using ANOVA analysis and a 2-tailed *t*-test, no significant improvement is seen in the DSS group over the No DSS group. However, due to the effect size being smaller than what was initially expected based on the results of Experiments I and II, there is insufficient design power to state this conclusively. Despite this, the data trends will nonetheless be reported but must be considered observational.

Table 4: Summary of mean results for mild ice concentration groups

$4/10^{th}$	Mean Con Change ($X/10^{th}$)	Max Con Change ($X/10^{th}$)	Mean Total Clearing (km^2)	Max Time Clear (s)	Cumulative Time Clear (s)	Mean Clearing to Distance	First Time Clear (s)
DSS Cadets	1.75	4.44	0.80	342	530	0.63	932
NoDSS Cadets	1.10	4.29	0.50	397	586	0.25	845
Seafarer	1.76	4.58	0.81	580	687	0.52	1043
T1	2.23	5.38	1.03	388	735	0.61	750
T2	2.10	5.08	0.96	420	547	0.70	1082

Table 5: Summary of mean results for severe ice concentration groups

7/10th	**Mean Con Change (X/10th)**	**Max Con Change (X/10th)**	**Mean Total Clearing (km^2)**	**Max Time Clear (s)**	**Cumulative Time Clear (s)**	**Mean Clearing to Distance**	**First Time Clear (s)**
DSS	1.76	5.02	0.81	311	461	0.49	998
No DSS	1.87	5.33	0.86	294	407	0.48	987
Seafarer	2.78	5.94	1.28	545	710	0.86	895
T1	2.48	6.57	1.13	400	567	0.95	1157
T2	2.95	6.82	1.35	618	743	1.13	950

A two-tailed *t*-test, assuming equal variance was then performed to check the statistical significance of each group compared. These are shown in Table 6 and Table 7. Each set of comparisons is seen in the left-hand column, and the *p*-value of the difference between these groups is shown. From this, it is clear that there is no statistical significance in the improvement of the DSS group over the No DSS group. Values that are not given an asterisk do not have 80% power. Values with an asterisk do have at least 80% power. In these tables, only those that have a *p*-value from their *t*-test greater than 0.05 have a statistically significant difference at 95% confidence. In the severe ice concentration group, only the Training I group performs significantly better than the DSS group in the Mean Concentration Change, Max Concentration Change, Mean Total Clearing, and Mean Clearing to Distance Ratio metrics. It should be noted that the DSS group performs significantly worse than the Training II group in the severe ice conditions group for 5 out of 7 metrics. Although observationally, the DSS group performs worse than Seafarers and Training I and Training II groups, for the most part, this is not statistically significant at 95% confidence.

Table 6: *t*-tests for mild ice conditions. Values marked with a '*' have >80% design power and values without have <80% design power.

t-test: 4/10th	**Mean Conc Change**	**Max Conc Change**	**Mean Total Clearing**	**Max Time Clear**	**Cumulative Time Clear**	**Mean Clearing to Distance**	**First Time Clear**
DSS to No DSS	0.101	0.841	0.097	0.748	0.738	0.027	0.483
DSS to Seafarer	0.975	0.860	0.987	0.090*	0.194	0.564	0.155
DSS to T1	0.314*	0.304*	0.313*	0.715	0.116	0.918	0.125
DSS to T2	0.318*	0.368*	0.336*	0.521*	0.889	0.719*	0.143

Table 7: *t*-tests for severe ice concentration groups. Values marked with a '*' have >80% design power and values without have <80% design power.

t-test: 7/10th	**Mean Conc Change**	**Max Conc Change**	**Mean Total Clearing**	**Max Time Clear**	**Cumulative Time Clear**	**Mean Clearing to Distance**	**First Time Clear**
DSS to No DSS	0.786	0.711	0.784	0.878	0.698	0.942	0.951
DSS to Seafarer	0.018	0.277	0.017	0.125*	0.152	0.053	0.611
DSS to T1	0.088*	0.080*	0.089*	0.338	0.414	0.022	0.380
DSS to T2	0.005*	0.050*	0.005*	0.007*	0.044	0.002*	0.799

4.2.1 Average Clearing

In examining the results for the Average Clearing metric half-normal plots for both the subplot effects and whole-plot effects can be seen in Figure 14. This was performed as a part of the significance check to determine which if any points were significant for this metric (Montgomery, 2013). For this metric, neither the influence of the DSS, nor the ice concentration factors were found to be statistically significant when $\alpha = 5\%$, meaning at 95% confidence. This indicates that there is functionally no improvement in ice management performance for the metric of average ice clearing when using the DSS.

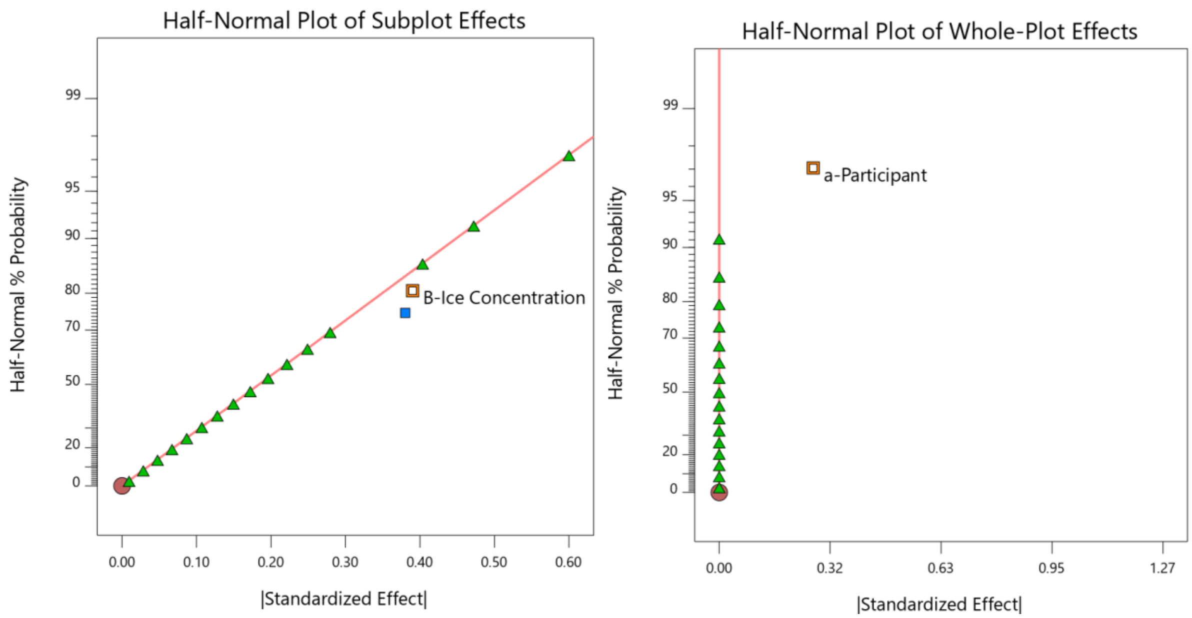


Figure 14: Half-Normal plot of Subplot effects (left) and Whole-Plot effects (right)

In order for this analysis to be completed, several assumptions about the results must be validated. This is done through an analysis of the residuals using the software Design Expert 11 for each metric. The first assumption states that for the analysis to take place as a full-factorial experiment, the results must be normally distributed. This is done by arranging the residuals on a normal plot. If the residuals are all distributed relatively close to the normal line, then the results can be considered normal. As can be seen in Figure 15, this assumption holds true as all residual points are distributed relatively close to the normal line in a linear fashion.

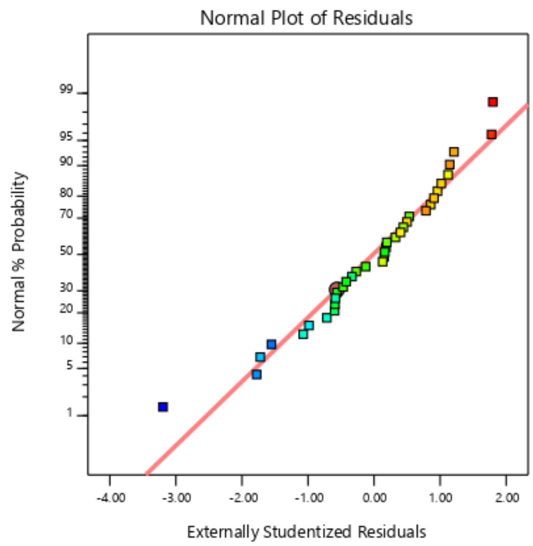


Figure 15: Normal plot of residuals. Since these are all close to the Normal line, the assumption of normality is considered valid.

The next assumption that must be validated is that of heterodascitiy, meaning that the two samples are assumed to have equal variance and are randomly scattered. In order to validate this, the residuals are plotted and must be found to be in between the upper and lower boundaries, depicted by horizontal lines in the plot. As this can be seen to be the case from Figure 16, this is indeed the case and indicates that the data are randomly scattered. This validates the assumption of equal variance.

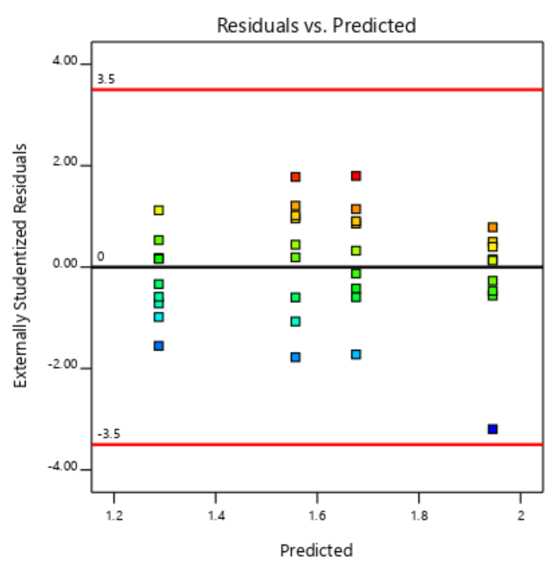


Figure 16: residuals vs. predicted plot. Since all residuals are between the upper and lower bounds of ±3.5, the assumptions of equal variance are considered valid.

The next assumption that must be validated is that of random scatter, that is, that no time-related variables are affecting the results. This is especially important in this experiment, as it is known that due to logistics issues with scheduling participants, and comparing different groups across different studies, the results are not truly random. As the analysis methods depended on assumptions of random scatter, it was important to ensure that the steps taken to maximise the randomness of the results as much as possible held true. In Experiment III, the same steps and instructions scripts were used as in Experiments I and II to ensure that they were as similar as possible. Also, as discussed, participants were effectively randomly assigned to different ice concentration groups. When checking that this assumption of randomness holds true, it can be seen in Figure 17 in the Residuals vs. Run Order plot that the residuals are distributed in between the upper and lower boundaries. This indicates that this assumption does indeed hold true.

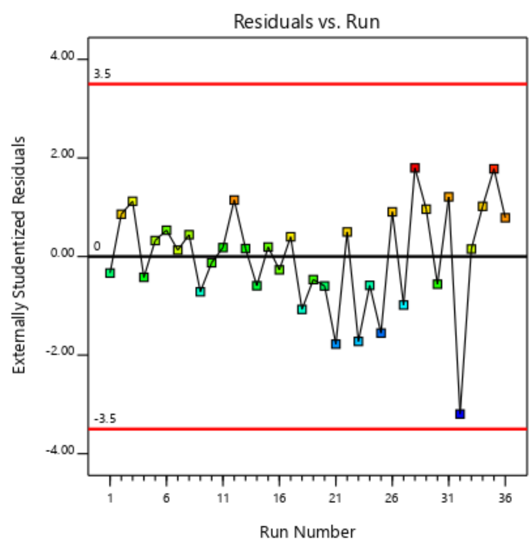


Figure 17: shows the residuals vs. run order plot. Since all points are between the upper and lower bounds, the assumption of independence of run can be considered valid.

Finally, it must be checked that no transformations are needed. The Box-Cox method can be used to induce normality or equal variance in the responses through the application of a transformation to the response variable (Montgomery, 2013). It can also improve the fit of the data to a particular model. For each response metric, the Box-Cox plot was generated using Design Expert II and was examined to see if any transformation was required to validate the assumptions of the model. For this metric, the Box-Cox plot is shown in Figure 18. The key value for this metric is λ, which represents the parameter of the transformation. The optimal λ value is represented in Figure 18 by the long vertical line, at the minima of the curve. The λ achieved is the shorter line which extends slightly below the x-axis. The confidence intervals of λ^+ and λ^- are represented by the thin vertical lines on either side of the optimal λ value. They are at the intersection points of the curve with the horizontal line denoting the log likelihood at 51.9042. If the λ achieved is within the confidence interval, then no transformation is required. As can be seen in Figure 18, the thin vertical line is comfortably within the confidence interval. This indicates that no transformation is required. If it fell outside of the lines, the software would recommend a transformation to apply to the model.

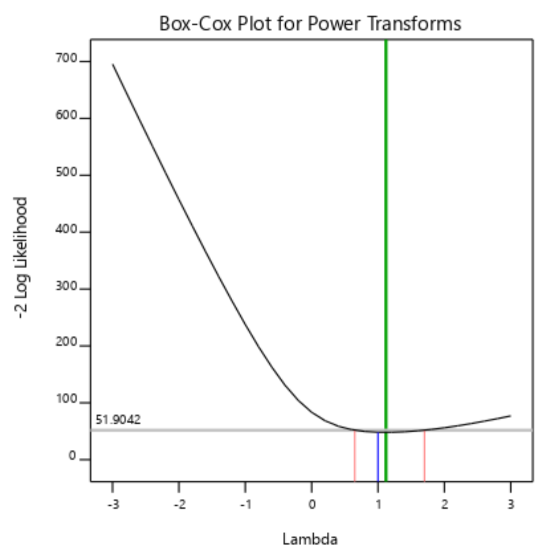


Figure 18: Box-Cox plot. This demonstrates that no transformation is suggested since the lambda falls within the bounds.

Below in Figure 20 and Figure 21, the results for the average clearing metric are shown, comparing the DSS group to the No DSS group, as well as the additional groups from Experiment I and Experiment II. An ANOVA table comparing the DSS and No DSS groups is presented in Table 8.

Table 8: ANOVA table for average clearing metric

Source	Term df	Error df	F-value	p-value	
Whole-plot	1	33.00	0.9549	0.3356	not significant
a-Participant	1	33.00	0.9549	0.3356	
Subplot	1	33.00	1.99	0.1678	not significant
b-Ice Concentration	1	33.00	1.99	0.1678	

The box plots shown below in Figure 20 and Figure 21 are an effective way to represent the results of this experiment, and compare them to the results of Experiments I and II. Box plots show the mean value, median value, upper and lower quartiles, and upper and lower extreme values. A description of the components of a box-plot can be seen in Figure 19.

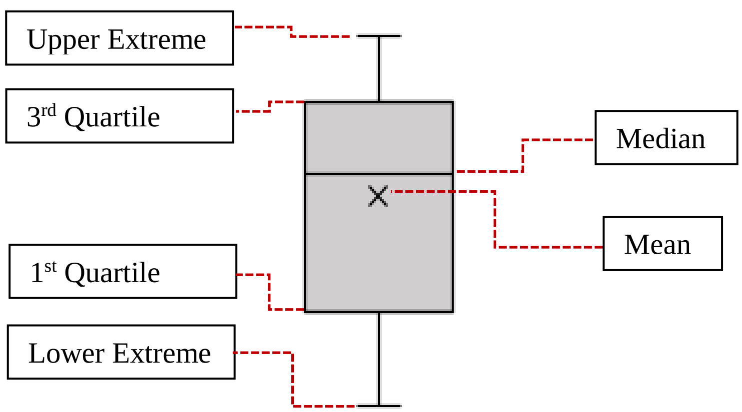


Figure 19: Description of Box Plot components *(Microsoft Corporation , n.d.)*

Figure *20* shows a box-plot of the results for the mild ice concentration groups. From this, it can be seen that there is a tendency for the DSS group to perform better than the No DSS group. This is shown by all components of the box plot: the extreme values, quartiles, mean, and median. The 3rd quartile and extreme high value are in fact higher than the best performance of the seafarer group, and the mean is similar, indicating that the inexperienced participants using the DSS in this group did find it advantageous. The trained cadets have the best performance of the group, as is discussed at length in Thistle and Veitch (2019). Figure *21* presents the results for the severe ice concentration groups. In this group, no appreciable difference is seen by the group using the DSS over the group without, and both are significantly worse than the seafarers and trained groups, as is discussed in Veitch et al., (2018) and Thistle and Veitch (2019).

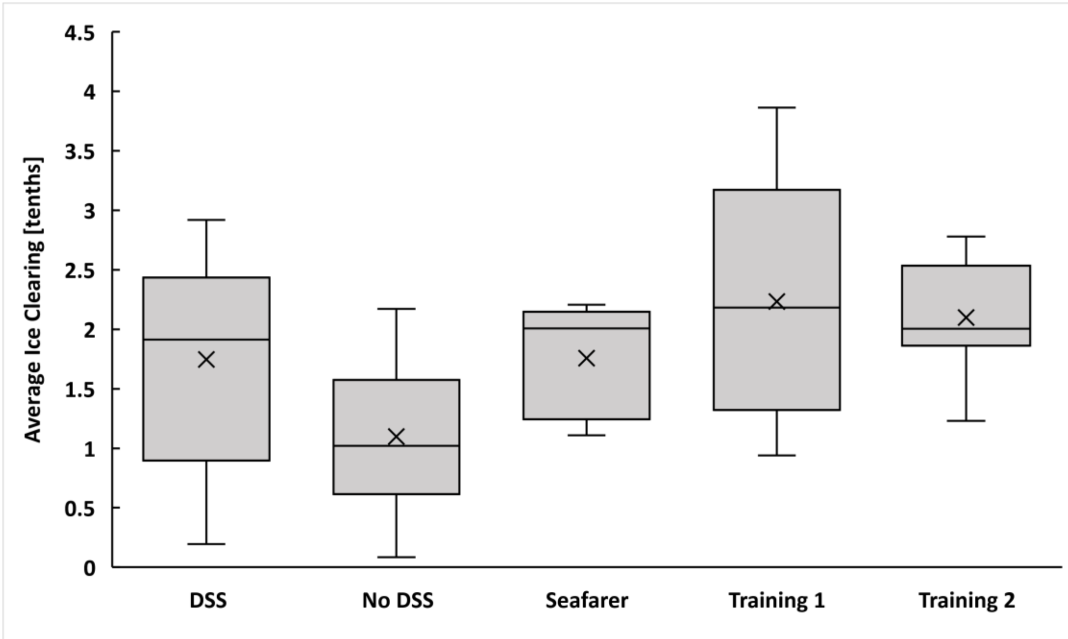


Figure 20: Mean concentration change, mild ice conditions

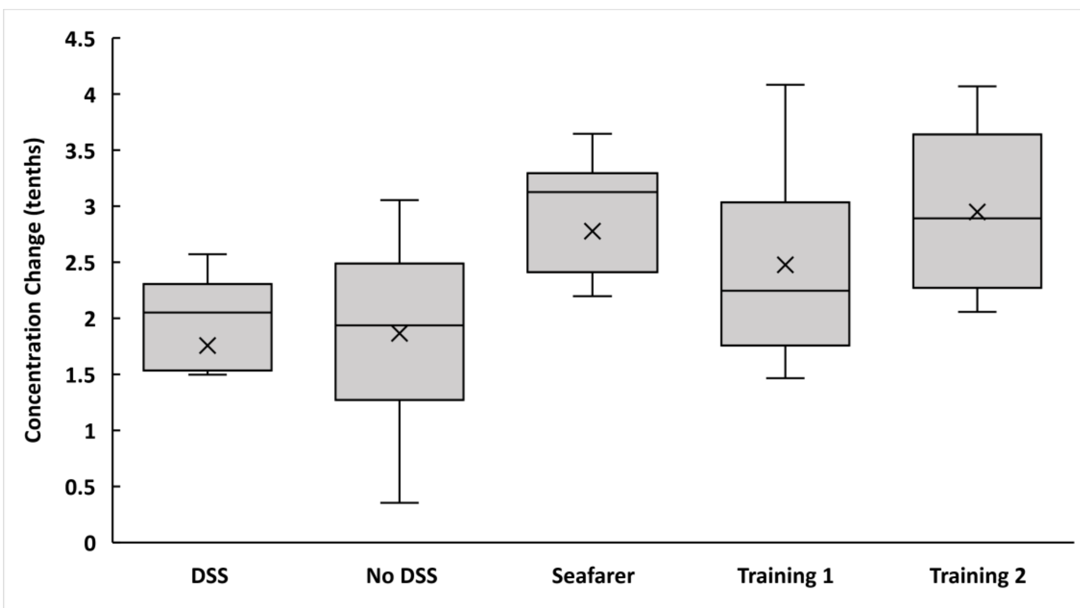


Figure 21: Mean concentration change, severe ice conditions

4.2.2 Maximum Clearing

There was no significant difference between the DSS and no DSS group (Table *9*). An analysis of residuals was performed and all assumptions were found to be valid, however in the Residuals vs. Run plot, there was a value on the borderline. The cause of this was a participant who had significant difficulties with the task and failed to clear any ice. The effect was no different than if no ice management was completed. In the exit interview, the participant stated that they forgot to use the side thrusters, so they were unable to complete the task in the way which was suggested by the DSS. There was no compelling reason to exclude the outlier however, so the data point was retained. For the sake of interest, if the data point is excluded, the ice concentration factor becomes significant, as can be seen in Table 10. No transformation was recommended for the model.

Table 9: ANOVA for maximum clearing metric with all data points included

Source	Term df	Error df	F-value	p-value	
Whole-plot	1	33.00	0.0201	0.8882	not significant
a-Participant	1	33.00	0.0201	0.8882	
Subplot	1	33.00	2.21	0.1468	not significant
b-Ice Concentration	1	33.00	2.21	0.1468	

Table 10: ANOVA table for maximum clearing with outlier removed

Source	Term df	Error df	F-value	p-value	
Whole-plot	1	32.00	0.2255	0.6381	not significant
a-Participant	1	32.00	0.2255	0.6381	
Subplot	1	32.00	5.60	0.0242	significant
b-Ice Concentration	1	32.00	5.60	0.0242	

Using an analysis of residuals similar to in 4.2.1, it was found the assumption of independence of run was violated in one case. The plots can be seen in Figure 22. There was also a small violation in the Residuals vs. Predicted plot for equal variance. No transformation was recommended however, and the plot of normality was acceptable. According to Montgomery, a small to moderate violation is not a concern when the violation does not refer to the assumption of normality (Montgomery, 2013).

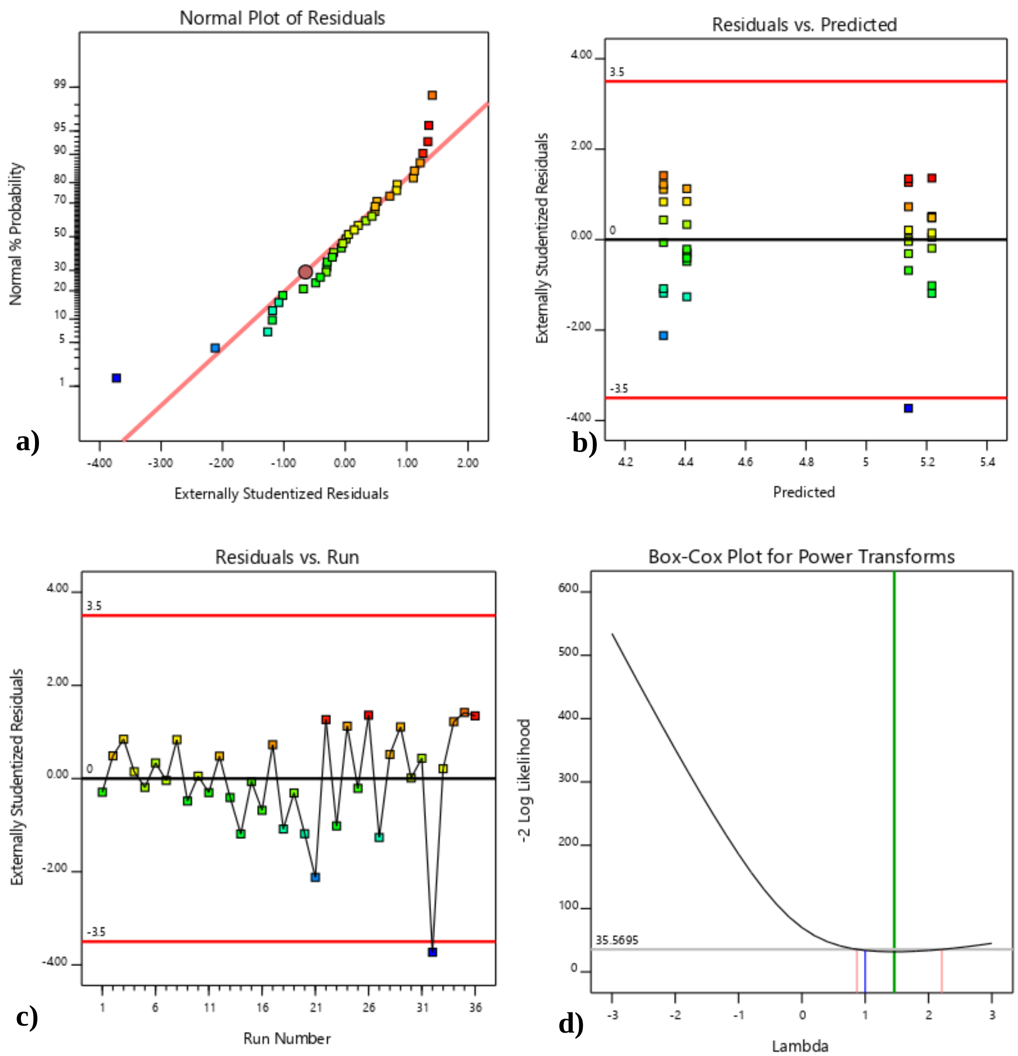


Figure 22: Example of diagnostic plots for residuals of Mean Concentration Change metric: a) shows the normal plot of residuals. Since these are all close to the Normal line, the assumption of normality is considered valid. b) shows the residuals vs. predicted plot. Since most of the residuals are between the upper and lower bounds, the assumptions of equal variance are considered valid. There is one outlier point, however it is not considered to be severe enough of a violation to reject the assumption. c) shows the residuals vs. run order plot. Since most points are between the upper and lower bounds, the assumption of independence of run can be considered valid. One point falls outside of the bounds, however this is not considered to be a severe enough violation to reject the assumption. d) shows the Box-Cox plot. This demonstrates that no transformation is suggested since the lambda falls within the bounds.

From the boxplots below in Figure 23 and Figure 24, it can be seen that despite the upper quartiles being better in the DSS group than the No DSS group, not much improvement can be observed. This trend holds true for the other metrics as well in both ice condition groups. When observing the mean value, the range is quite small, from 4.4/10th at the low end, to 5.3/10th at the high end in the mild ice condition group. In the severe ice concentration group, this range is slightly larger, from 5.0/10th to 6.6/10th, but it is still not dramatic. In observing the results, many of the participants were able to clear a large area, but only briefly and not with any consistency. This would present itself as the maximum clearing of a run.

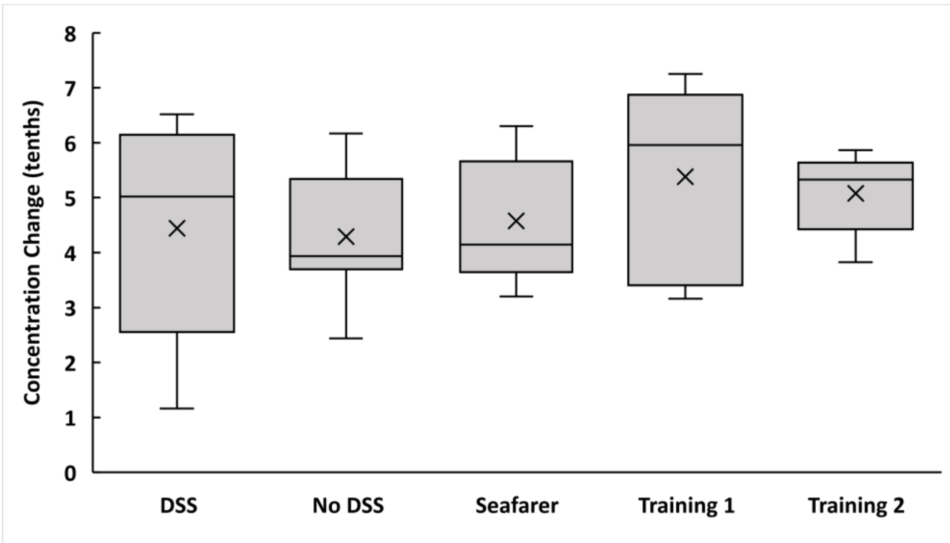


Figure 23: Maximum clearing, mild ice conditions

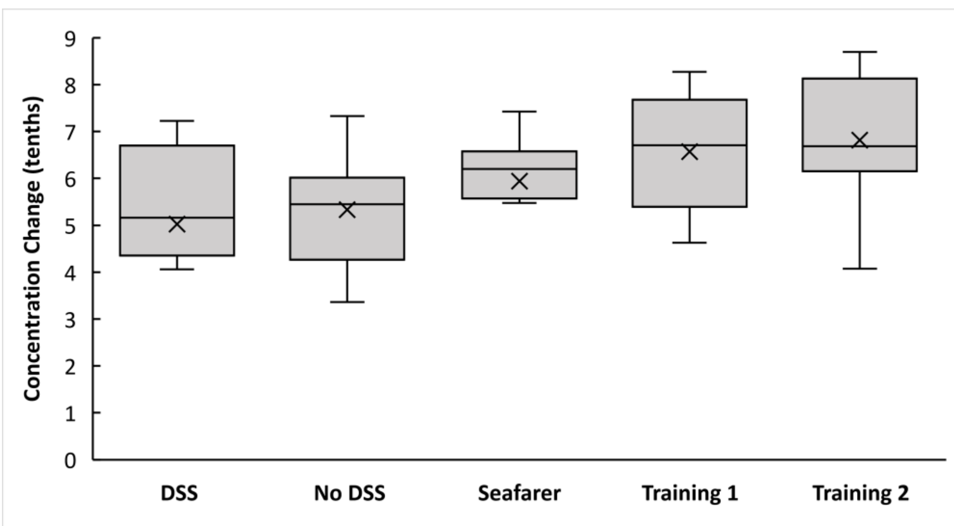


Figure 24: Maximum clearing, severe ice conditions

4.2.3 Total Clearing

The plots of residuals can be seen in Figure 25. From this it can be seen that all assumptions are accepted, however one outlier point lands on the lower bound of b) and c). This is not considered to be a severe violation however, so the assumptions are accepted.

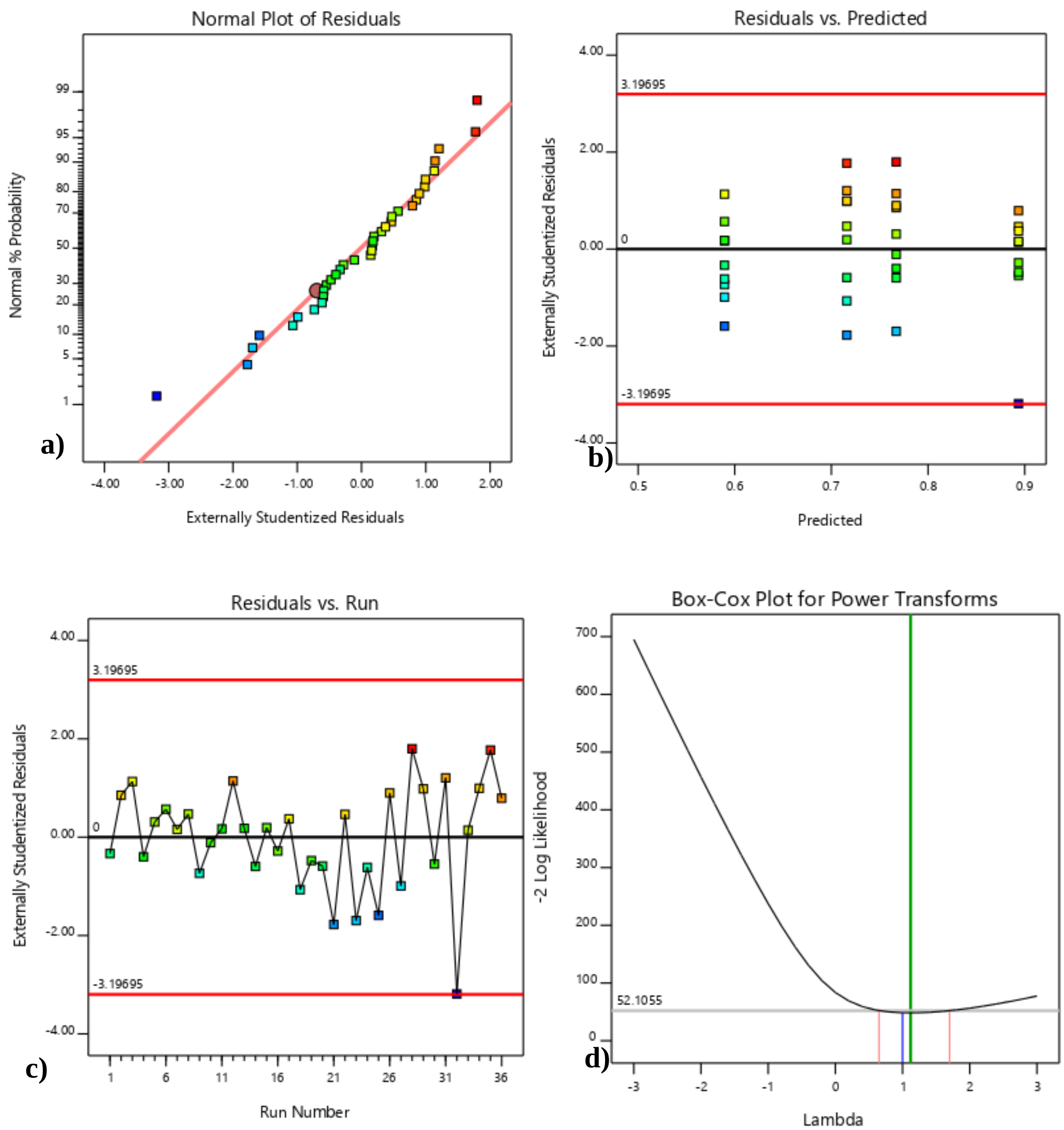


Figure 25: Example of diagnostic plots for residuals of Total Clearing metric: a) shows the normal plot of residuals. Since these are all close to the Normal line, the assumption of normality is considered valid. b) shows the residuals vs. predicted plot. Since most of the residuals are between the upper and lower bounds, the assumptions of equal variance are considered valid. There is one outlier point on the lower boundary, however it is not considered to be severe enough of a violation to reject the assumption. c) shows the residuals vs. run order plot. One point falls on the lower, however this is not considered to be a severe enough violation to reject the assumption. d) shows the Box-Cox plot. This demonstrates that no transformation is suggested since the lambda falls within the bounds.

As can be seen from the ANOVA table, Table 11, comparing the DSS and No DSS groups, there is no significant difference in performance due to the *p*-value being > 0.05. There is also no significant difference between the ice concentration groups, which indicates that ice concentration is not a factor in the ice management performance for this metric.

Table 11: ANOVA for Total Clearing metric

Source	Term df	Error df	F-value	p-value	
Whole-plot	1	33.00	0.9982	0.3250	not significant
a-Participant	1	33.00	0.9982	0.3250	
Subplot	1	33.00	1.97	0.1698	not significant
b-Ice Concentration	1	33.00	1.97	0.1698	

The Total Clearing metric tabulates the total quantity of ice cleared throughout a run, measured in units of area. An increase in performance for the DSS group over the No DSS group can be observed in the mild ice conditions group (Figure 26). The mean value is in fact similar to that of the seafarer group, and the upper quartile and extreme high value is higher than this group. The Training I and Training II groups are higher than all other groups, with the Training I group having the best performance overall. Figure 27 displays the results for the same metric, but for the severe ice concentration group. Although less variability is observed in the DSS group, the performance can be observed to be similar to that of the No DSS group, and is significantly smaller than the seafarer and training groups.

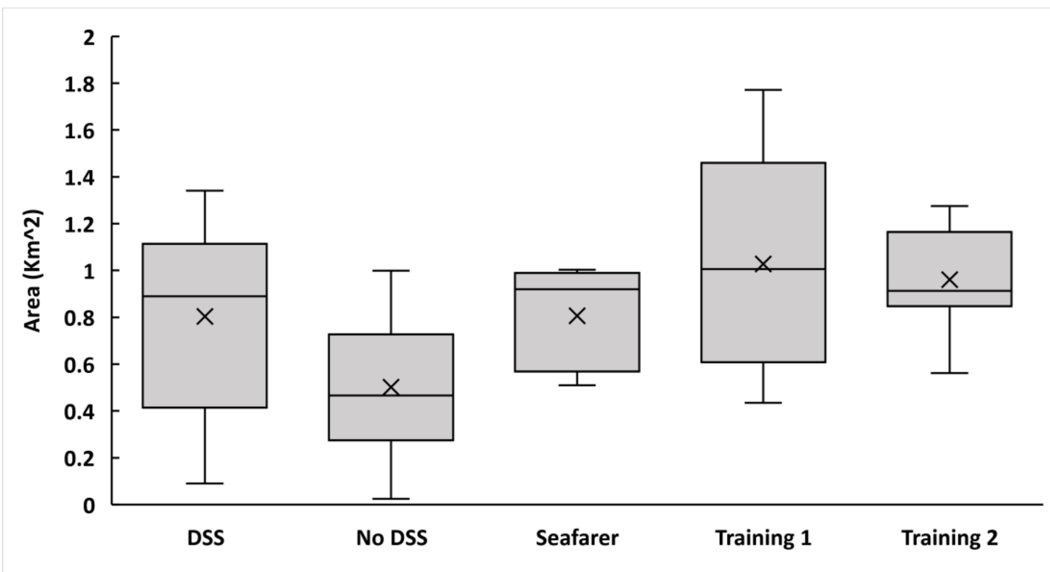


Figure 26: Mean total clearing, mild ice conditions

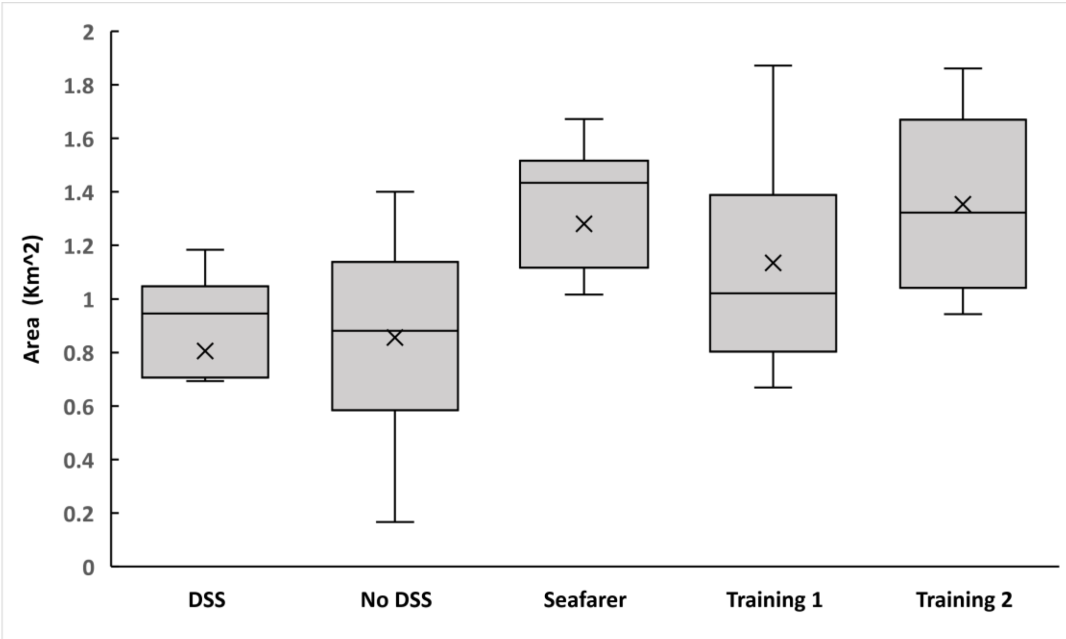


Figure 27: Mean total clearing, severe ice condition

4.2.4 Clearing to Distance ratio

In examining the plots of residuals in Figure 28, it can be seen that all assumptions can be considered valid, and that no transformation is required.

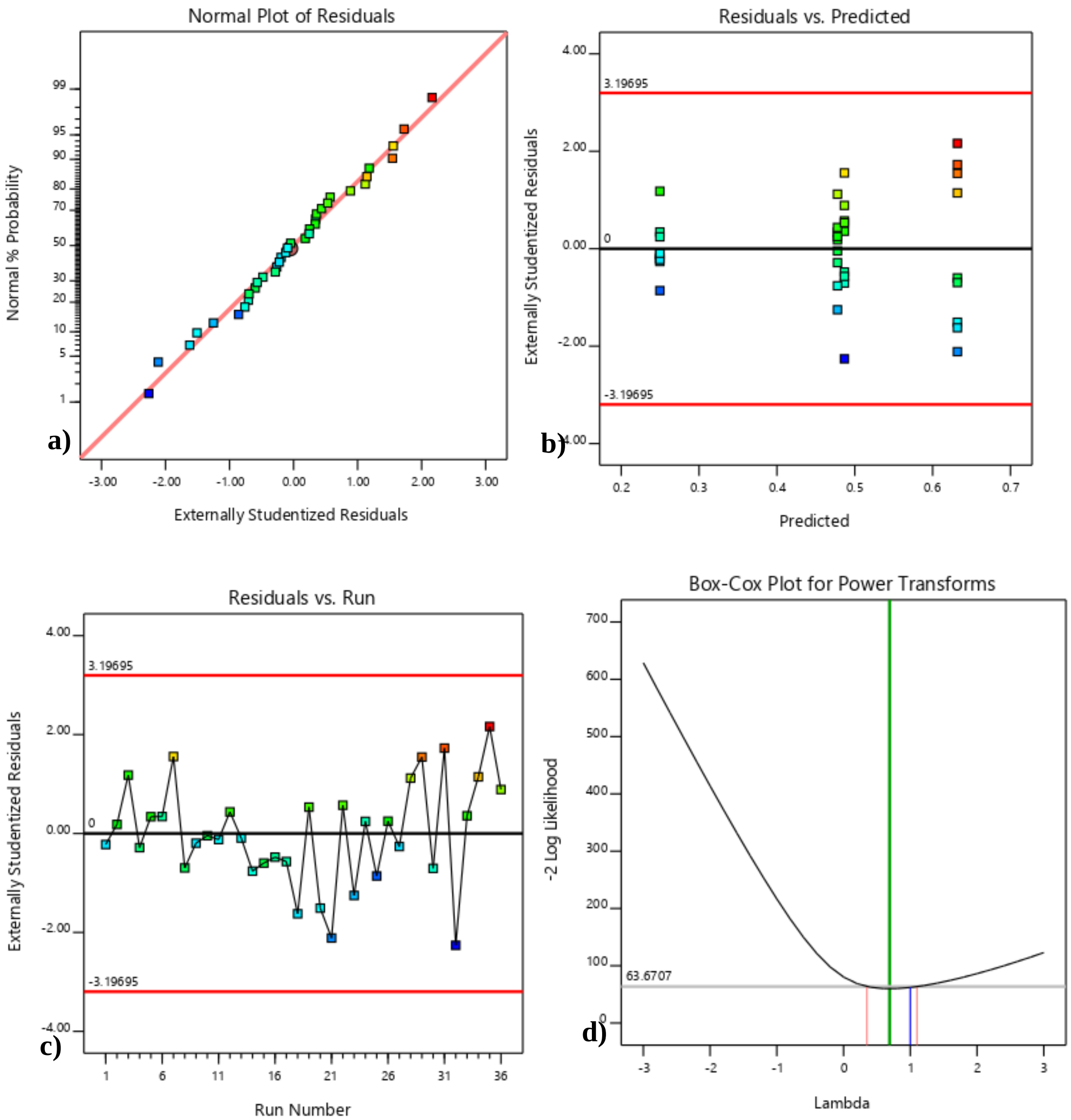


Figure 28: Diagnostic plots for residuals of Mean Concentration Change metric:. a) shows the normal plot of residuals. Since these are all close to the Normal line, the assumption of normality is considered valid. b) shows the residuals vs. predicted plot. Since all residuals are between the upper and lower bounds, the assumptions of equal variance are considered valid.. c) shows the residuals vs. run order plot. Since all points are between the upper and lower bounds, the assumption of independence of run can be considered valid. d) shows the Box-Cox plot. This demonstrates that no transformation is suggested since the lambda falls within the bounds.

For the Clearing to Distance ratio when comparing the DSS and No DSS groups, it can be seen in Table 12 that although the participant *p*-value approaches 0.05, it is still greater than this value, and thus is not significant. The ice concentration is also not a significant term.

Table 12: ANOVA for Clearing to Distance ratio

Source	**Term df**	**Error df**	**F-value**	**p-value**	
Whole-plot	1	33.00	3.62	0.0659	not significant
a-Participant	1	33.00	3.62	0.0659	
Subplot	1	33.00	0.1632	0.6889	not significant
b-Ice Concentration	1	33.00	0.1632	0.6889	

When examining the results of the mild ice concentration group, it can be observed in Figure 29 that the DSS group is superior to the No DSS group for all components of the box-plot. The mean, upper quartile, and upper extreme values are also higher than both the Seafarer and Training I groups, however Training II has a higher upper extreme value and mean. The median for the DSS group is lower than all other groups, with the exception of the No DSS group. When observing the results for this metric, four participants in the DSS group performed well. They were effective at clearing ice, but more importantly for this metric, they positioned their vessels to block ice using leeway and thus did not travel far. The shorter distance resulted in an increased clearing to distance ratio due to the small denominator in the calculation. In Figure 30, it can be seen that although the extreme value for the DSS group is higher than the No DSS group, the rest of the result cannot be observed to be much improved from the No DSS group. Additionally, the DSS group is lower than the Seafarer and Training metrics.

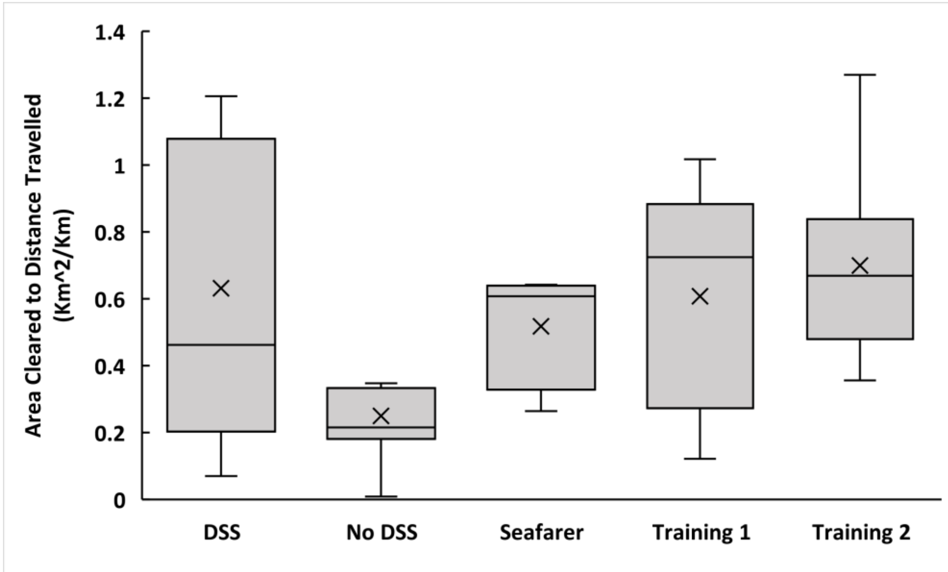


Figure 29: Mean clearing to distance ratio, mild ice conditions

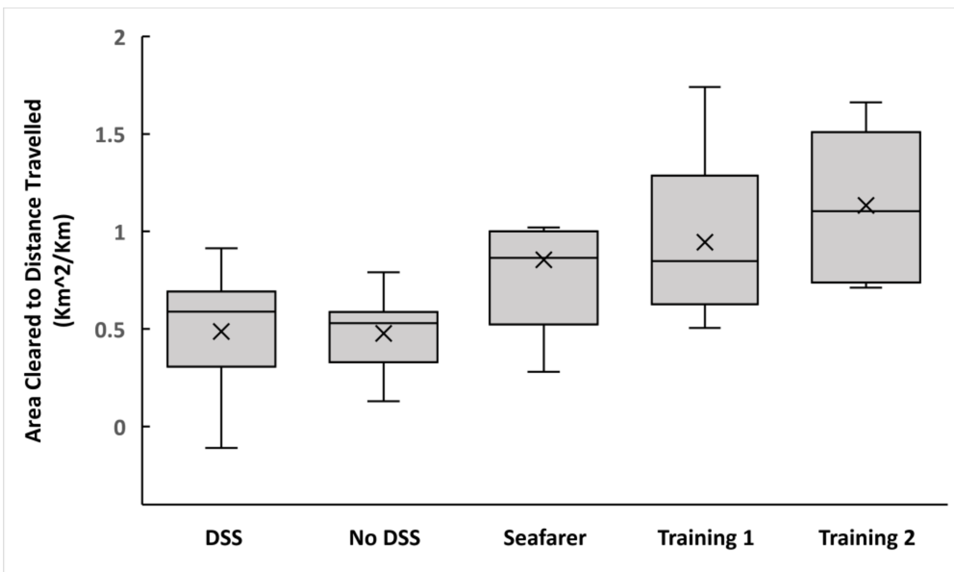


Figure 30: Mean clearing to distance ratio, severe ice conditions

4.2.5 Cumulative Lifeboat Launch Time

An initial analysis of the Box-Cox plot showed that a square root transformation was recommended to improve the validity of the assumptions (Figure 31). Although no clear violation were observed, this transformation was applied.

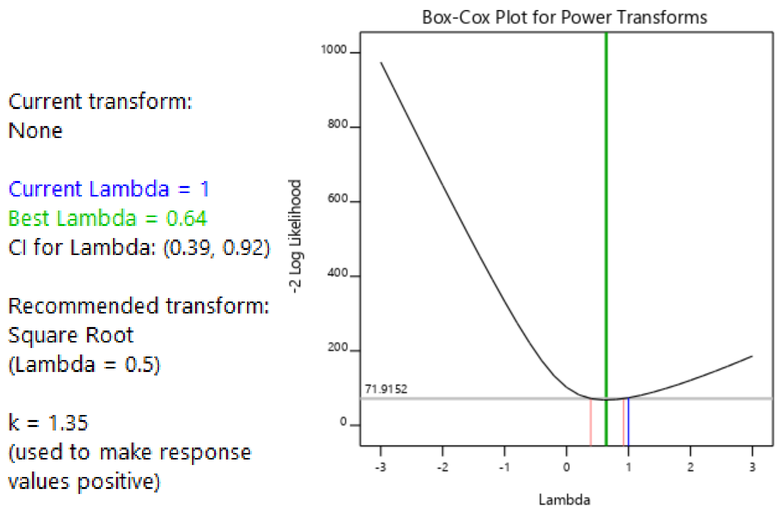


Figure 31: Box-Cox plot suggesting square root transformation. It can be seen that the lambda value falls outside of bounds.

After applying this transformation, it can be seen in Figure 32 that all assumptions can be considered to be valid, and that no additional transformations are required.

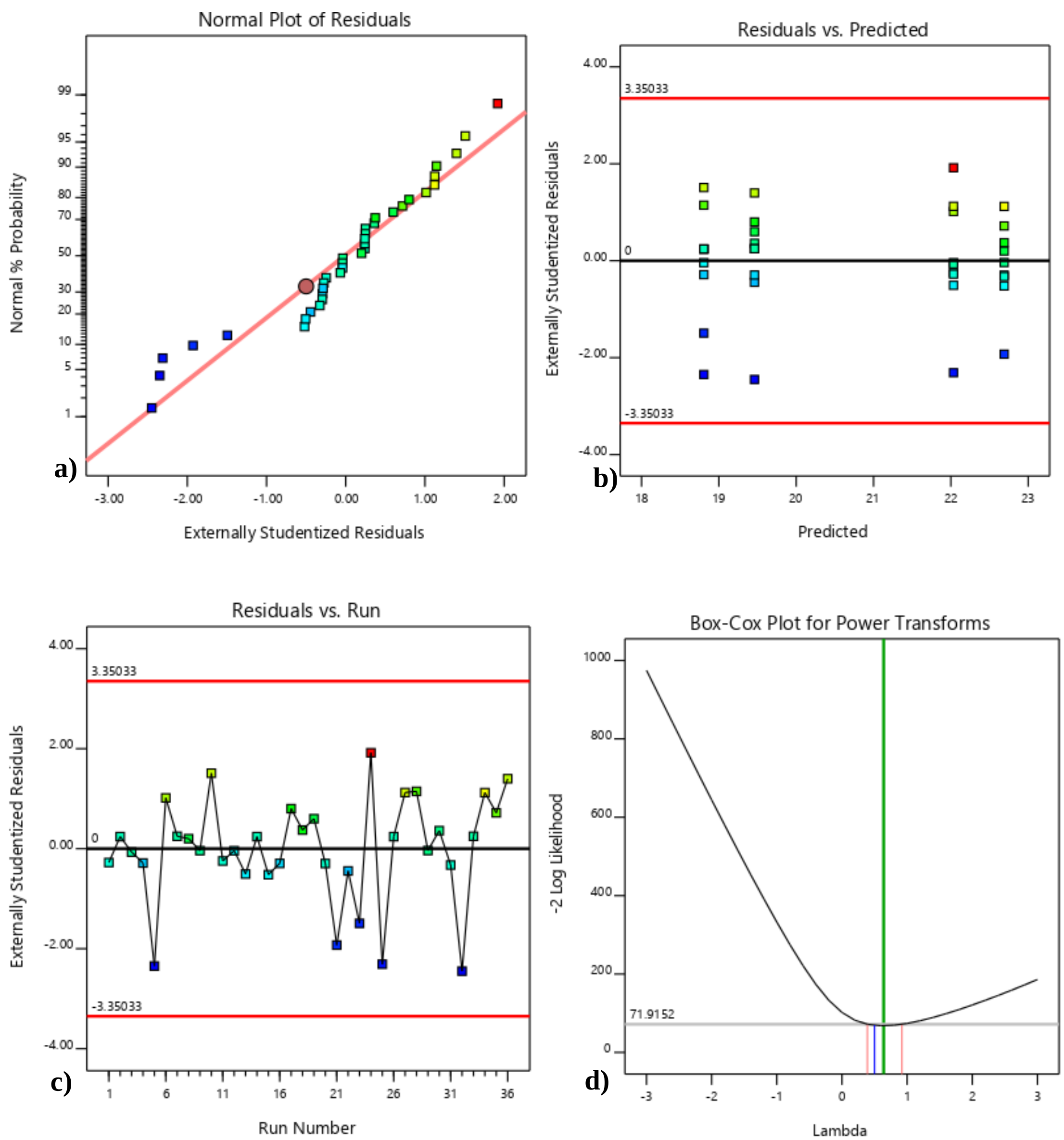


Figure 32: Example of diagnostic plots for residuals of Maximum Lifeboat Launch Time metric:. a) shows the normal plot of residuals. Since these are all close to the Normal line, the assumption of normality is considered valid. b) shows the residuals vs. predicted plot. Since all residuals are between the upper and lower bounds, the assumptions of equal variance are considered valid.. c) shows the residuals vs. run order plot. Since all points are between the upper and lower bounds, the assumption of independence of run can be considered valid. d) shows the Box-Cox plot after the square root transformation was applied.

Table 13 shows the ANOVA analysis for the Cumulative Launch time when comparing the DSS and No DSS groups. From this, it can be seen that with *p*-values greater than 0.05, the differences between the two groups are not statistically significant. Additionally, there is no significant differences between the ice concentration groups, indicating that this result is not dependent on ice concentration.

Table 13: ANOVA for Cumulative Lifeboat Launch Time

Source	**Term df**	**Error df**	**F-value**	**p-value**	
Whole-plot	1	16.00	0.0001	0.9918	not significant
a-Participant	1	16.00	0.0001	0.9918	
Subplot	1	17.00	1.43	0.2482	not significant
b-Ice Concentration	1	17.00	1.43	0.2482	

Observationally, it can be seen from both Figure 33 and Figure 34 that the DSS group does not appear to have any advantage over the No DSS group, and falls short of the Seafarer and Training groups. This holds true for both low and high ice concentration groups. There is significant variability in the results, with some participants performing quite well, and others not managing to clear the lifeboat launch zone for long, if at all. This can be observed in the extreme values presented by the whiskers.

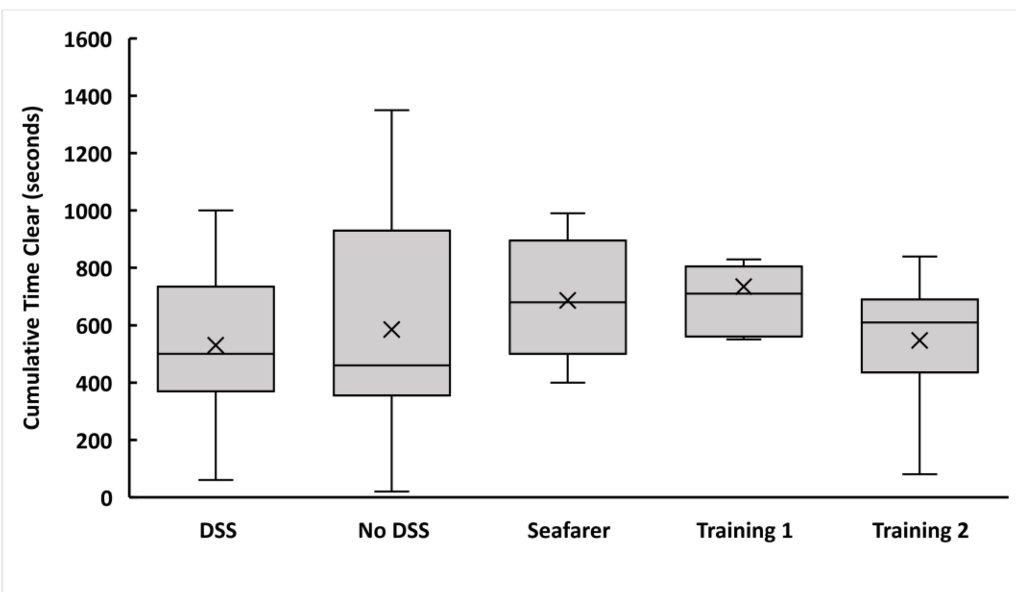


Figure 33: Cumulative time clear, mild ice conditions

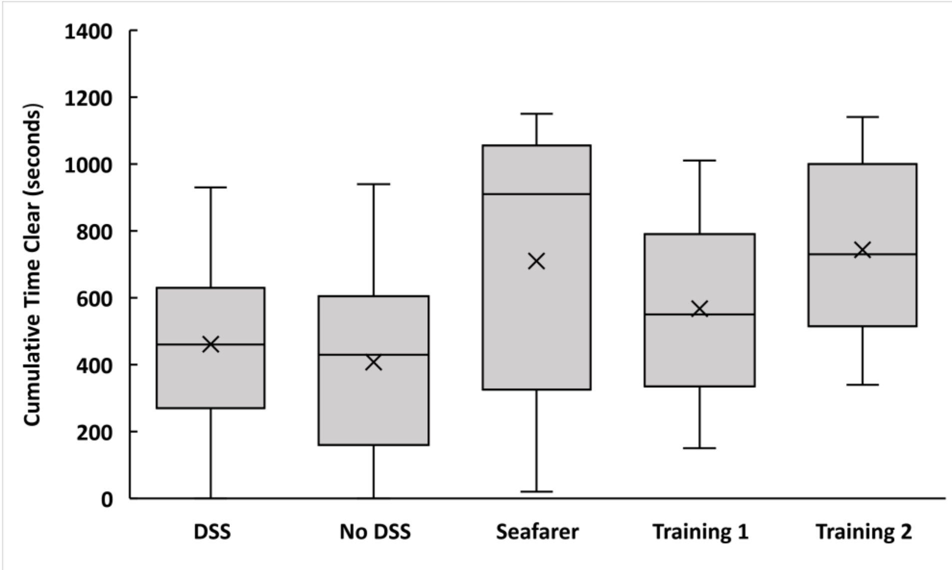


Figure 34: Cumulative time clear, severe ice conditions

4.2.6 Maximum Lifeboat Launch Time

Initial analysis of the normal plot of residuals showed that there were violations in the assumption of normality, as can be seen in Figure 35. From this, it is clear that the residuals do not appear to follow a linear trend in the plot. The Box-Cox plot, shown in Figure 36 recommended that a square root transformation be applied to induce normality, a technique generally used in this type of statistical analysis (Montgomery, 2013).

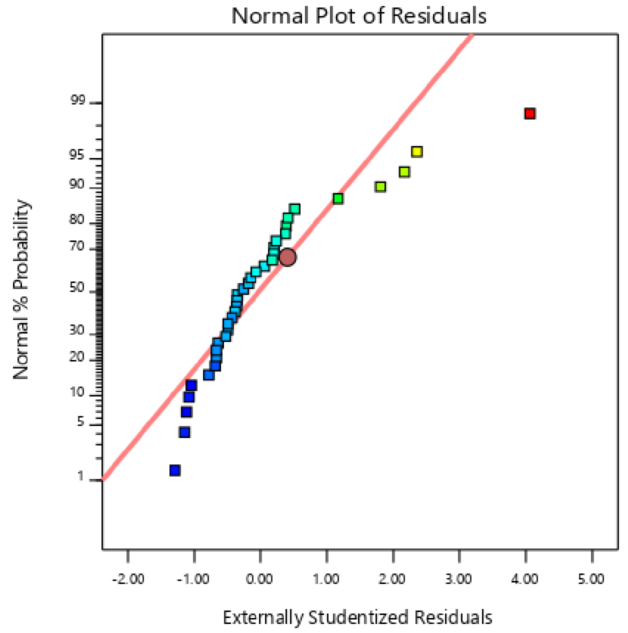


Figure 35: Initial normal plot of residuals showing violation of the assumption of normality before transformation is applied

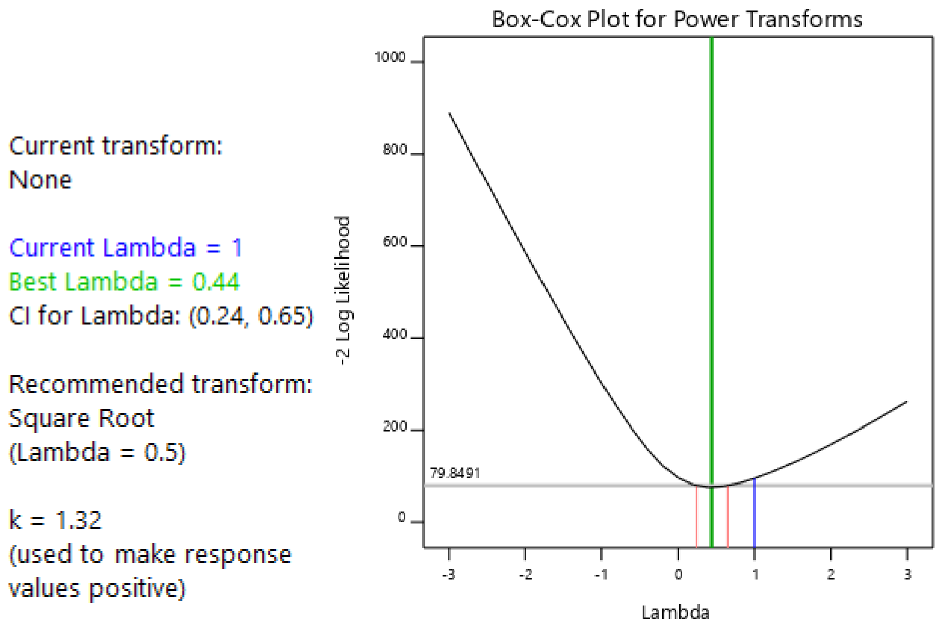


Figure 36: Box-Cox plot showing assumption violation and recommending Square Root transformation

Following the application of this transformation, it can be seen from Figure 37 that all assumptions have been validated. The most clear difference is in comparing Figure 37a to Figure 35, where it can be seen that normality has been induced.

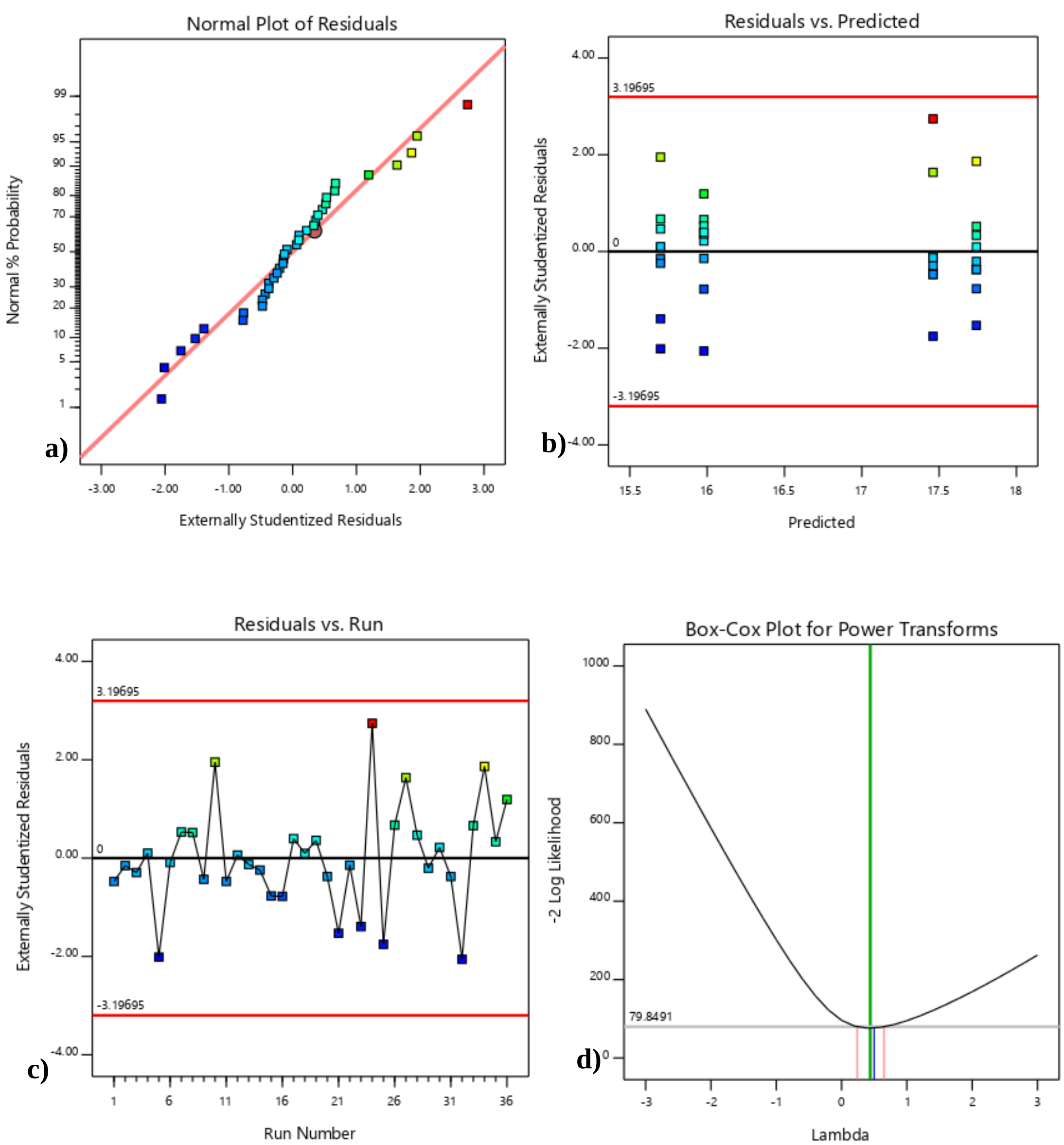


Figure 37: Example of diagnostic plots for residuals of Maximum Lifeboat Launch Time metric:. a) shows the normal plot of residuals. Since these are all close to the Normal line, the assumption of normality is considered valid. b) shows the residuals vs. predicted plot. Since all residuals are between the upper and lower bounds, the assumptions of equal variance are considered valid.. c) shows the residuals vs. run order plot. Since all points are between the upper and lower bounds, the assumption of independence of run can be considered valid. d) shows the Box-Cox plot after the square root transformation was applied.

Table 14 shows the results of the ANOVA analysis for the Maximum Lifeboat Launch Time metric. This metric is the maximum time interval in which the lifeboat launch zone is clear. From the table, it can be seen that the DSS group is not statistically different than the No DSS group, since the *p*-value is greater than 0.05. Additionally, it can be seen that the ice concentration does not result in significant results.

Table 14: ANOVA for Maximum Lifeboat Time Clear

Source	Term df	Error df	F-value	p-value	
Whole-plot	1	33.00	0.0114	0.9158	not significant
a-Participant	1	33.00	0.0114	0.9158	
Subplot	1	33.00	0.4488	0.5076	not significant
b-Ice Concentration	1	33.00	0.4488	0.5076	

In Figure 38, the results can be observed for this metric in the mild ice concentration group. It can be seen that the DSS group has a lower mean, lower upper extreme value, and lower upper quartile than the No DSS group. It is also lower than the Seafarer and Training groups. It can also be observed that these differences are slight, and that the mean and median values are close to those of the No DSS group, contributing to the lack of significance. Figure 39 provides the result for the severe ice concentration group. Some advantage can be observed in the DSS group, however it is nowhere near as advantageous as in the Seafarer and Training groups.

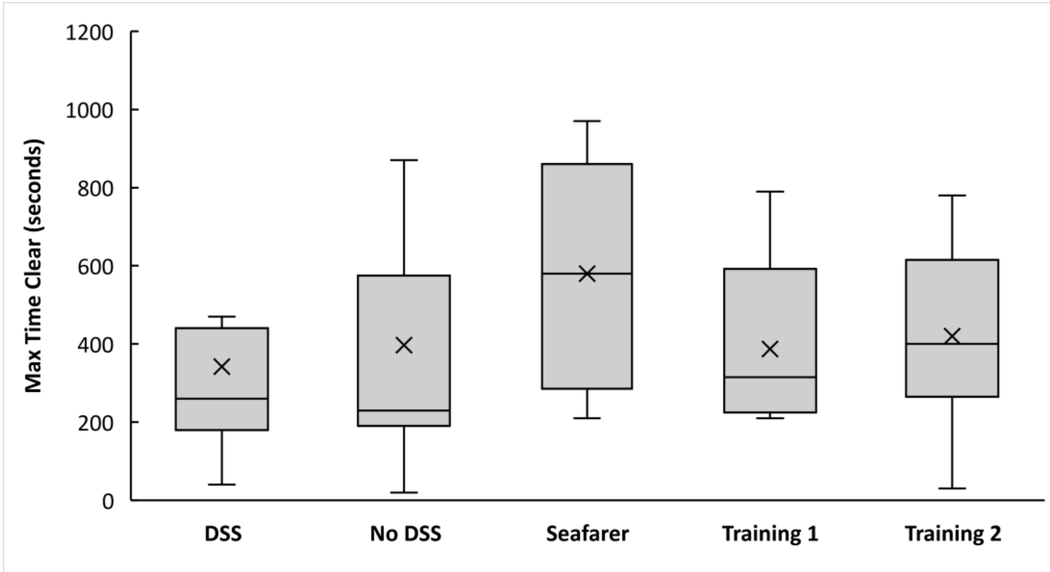


Figure 38: Maximum time clear, mild ice conditions

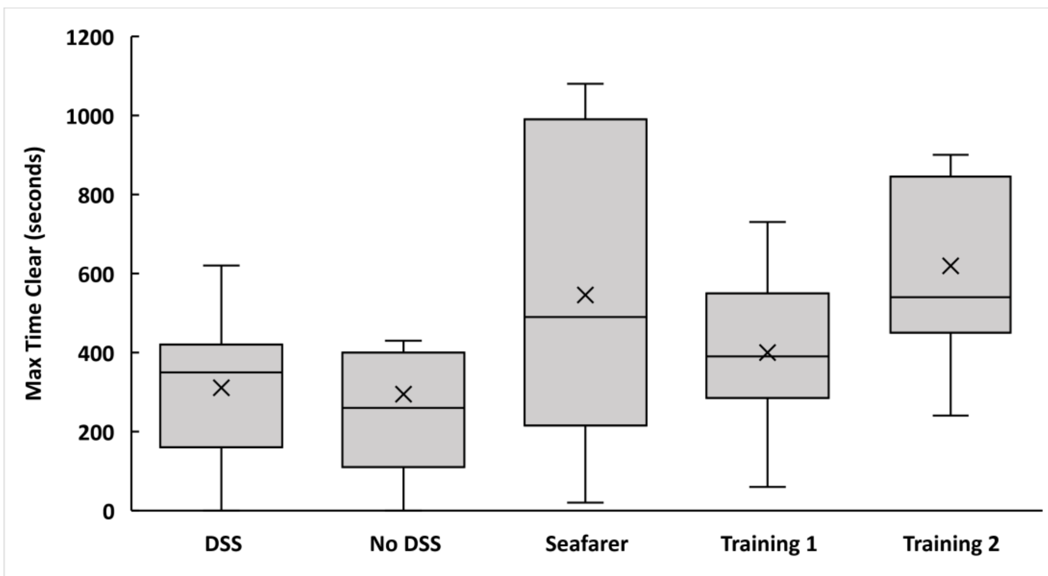


Figure 39: Maximum time clear, severe ice conditions

4.2.7 First Time Clear

Initial analysis of the Box-Cox plot, shown in Figure 41 recommended that a log transformation be applied to improve the model assumptions. In examining the plots of residuals, it can be seen from Figure 40 that although likely acceptable, the assumption of normality could be improved. Improving normality through transformations is a common technique for an analysis of this type (Montgomery, 2013).

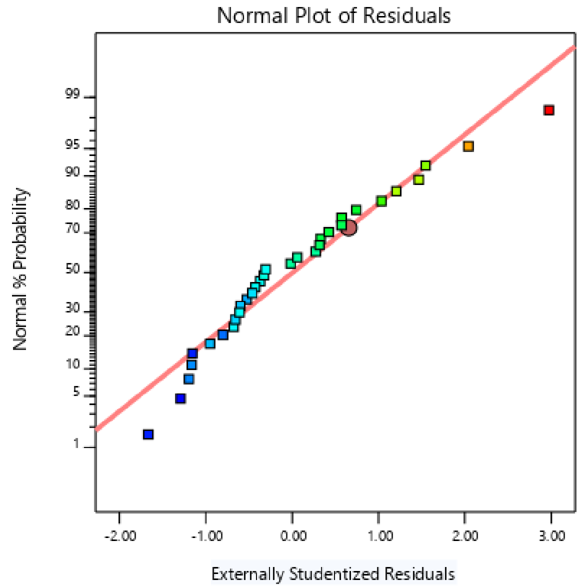


Figure 40: Normal plot of residuals before transformation is applied

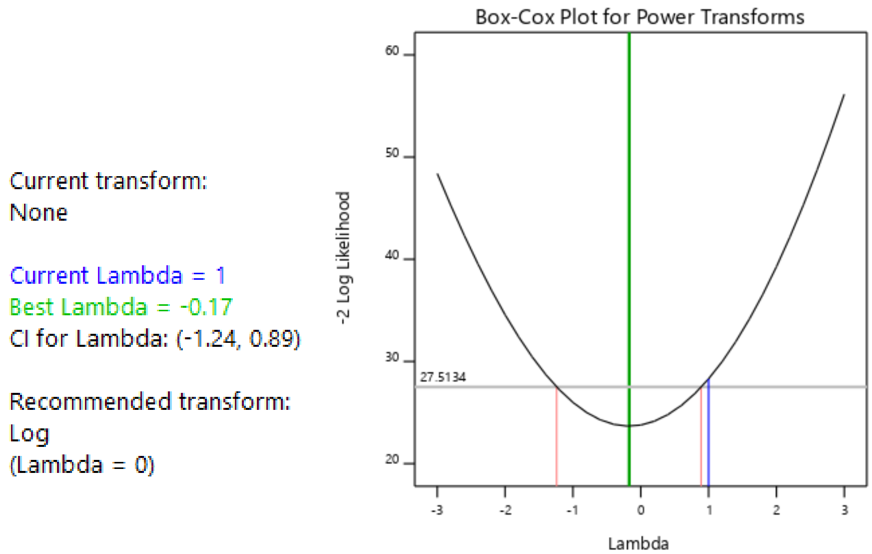


Figure 41: Box-Cox plot showing that a Log transformation is recommended

After applying the transformation, the plots of residuals are re-examined. From Figure 42a, it can be seen that the normality has been much improved over Figure 40, since the residuals are much closer to the normal line. All other residuals in Figure 42 can be seen to validate the assumptions of the model, including the new Box-Cox plot which shows that no further transformation is required.

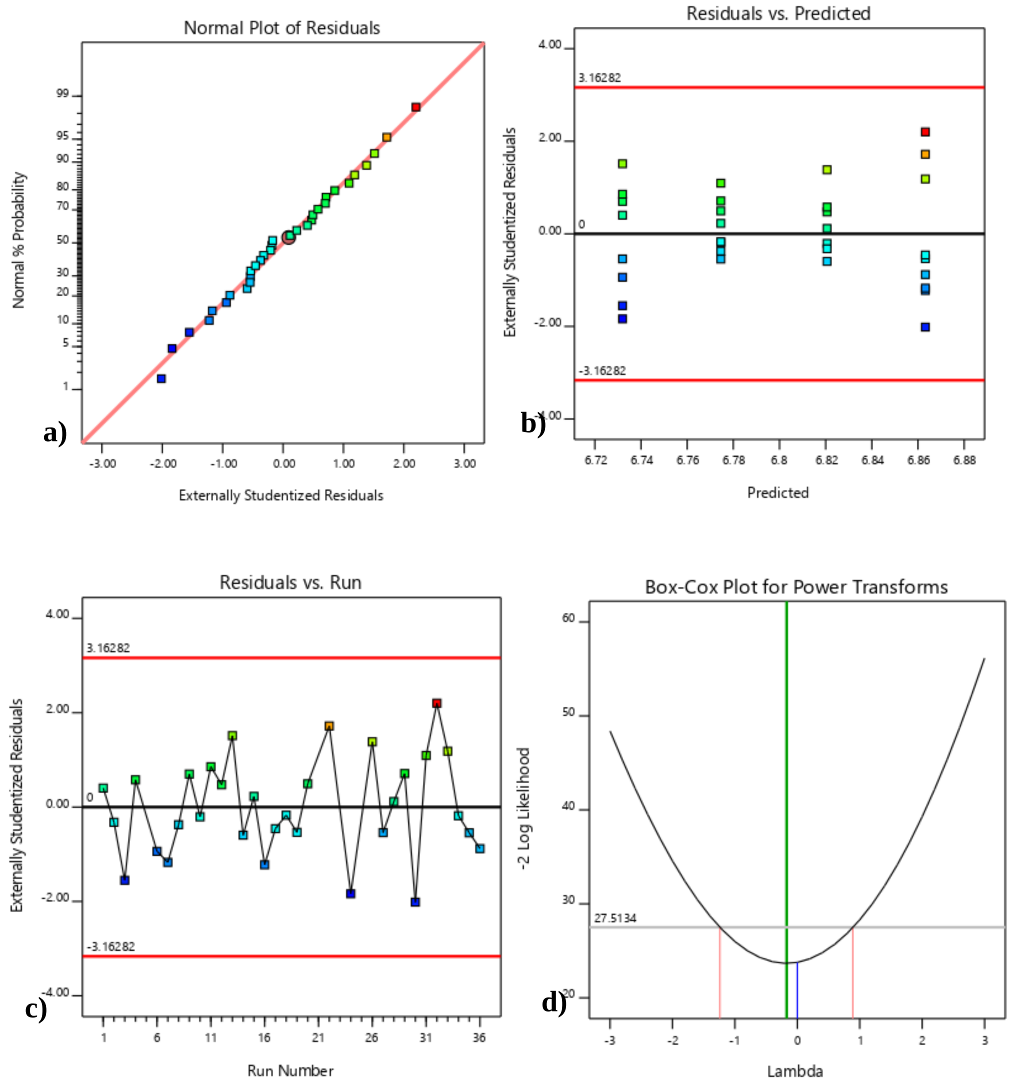


Figure 42: Diagnostic plots for residuals of Maximum Lifeboat Launch Time metric:. a) shows the normal plot of residuals. Since these are all close to the Normal line, the assumption of normality is considered valid. b) shows the residuals vs. predicted plot. Since all residuals are between the upper and lower bounds, the assumptions of equal variance are considered valid.. c) shows the residuals vs. run order plot. Since all points are between the upper and lower bounds, the assumption of independence of run can be considered valid. d) shows the Box-Cox plot after the square root transformation was applied.

In the First Time Clear metric, it can be observed from Table 15 that no significant difference exists between the DSS and No DSS groups. Furthermore, no significant difference can be found in between the ice concentration groups. This is due to the *p*-value being greater than 0.05.

Table 15: ANOVA for First Time Clear

Source	Term df	Error df	F-value	p-value	
Whole-plot	1	29.00	0.1406	0.7104	not significant
a-Participant	1	29.00	0.1406	0.7104	
Subplot	1	29.00	0.6112	0.4407	not significant
b-Ice Concentration	1	29.00	0.6112	0.4407	

In observing the box-plots for the mild ice concentration group in Figure 43, the lack of significant advantage in the DSS group is clear. This metric calculates the earliest time that the lifeboat is able to launch, so a lower value is considered better. The seafarer group and Training 2 groups seem to perform more poorly in this metric, however the difference is not dramatic when extreme values are discounted. Figure 44 provides the same result for the severe ice concentration metric. The difference in means for this metric are not significant between any of the groups, however it can be seen that the DSS group has both one of the best, and the worst results when extreme values are examined.

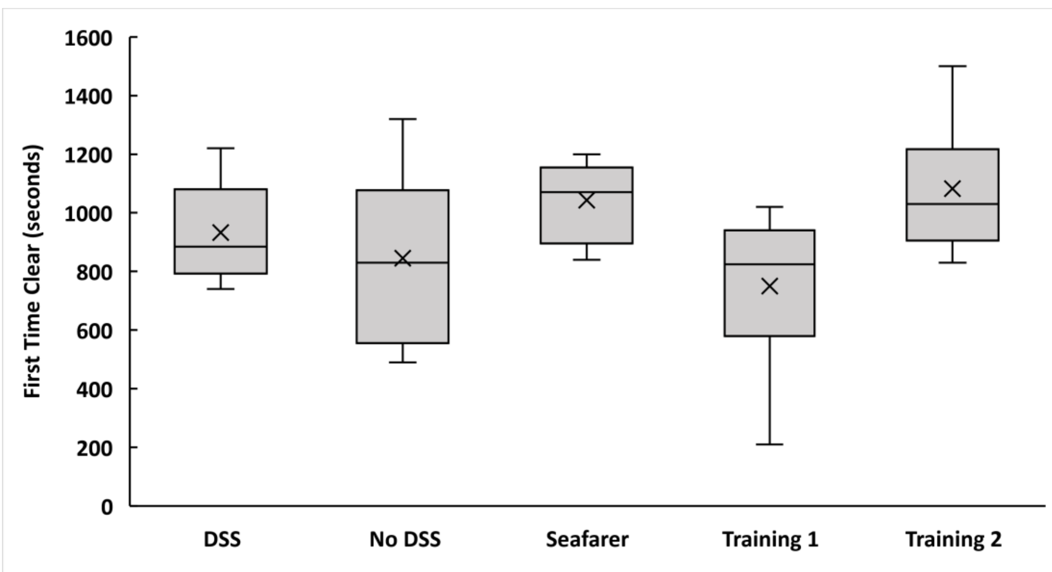


Figure 43: First time clear, mild ice conditions

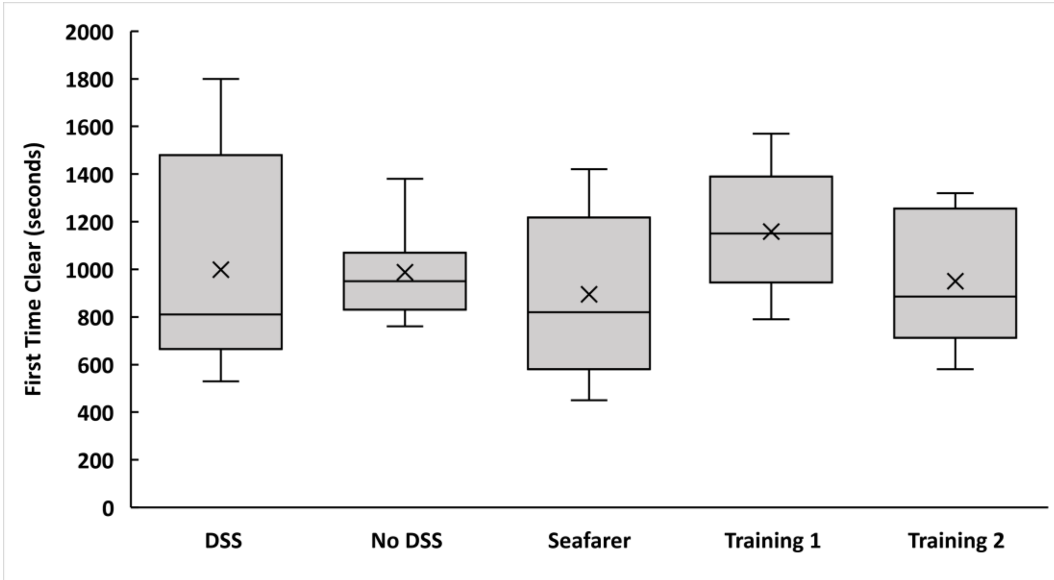


Figure 44: First time clear, severe ice conditions

4.3 Strategy Heat Maps

Despite not finding significant numerical results, it is possible to make some conclusions about the group behaviour by examining heat maps. Heat maps overlay the coordinates of each participant's own-ship in a group. Each second in which an own-ship is found in a pixel is coloured in, and the more seconds an own-ship is in the pixel, the lighter a colour it will be assigned. This means that in a group of 9 participants, for example, mild ice concentration, DSS group, a path in which multiple participants follow will receive a "hotter" value, showing as whiter coloured pixels. This provides an excellent way of observing the group behaviour, and general strategy. A sample of this can be seen in Figure 45, where the starting position of the own-ship is shown in relation to the FPSO. The position heat map can be seen overlaid here, with the legend displayed as well. For clarity, the heat maps will be further cropped to focus on the FPSO and the adjacent regions. The starting position is the same in all runs, but will not be shown, other than in Figure 45.

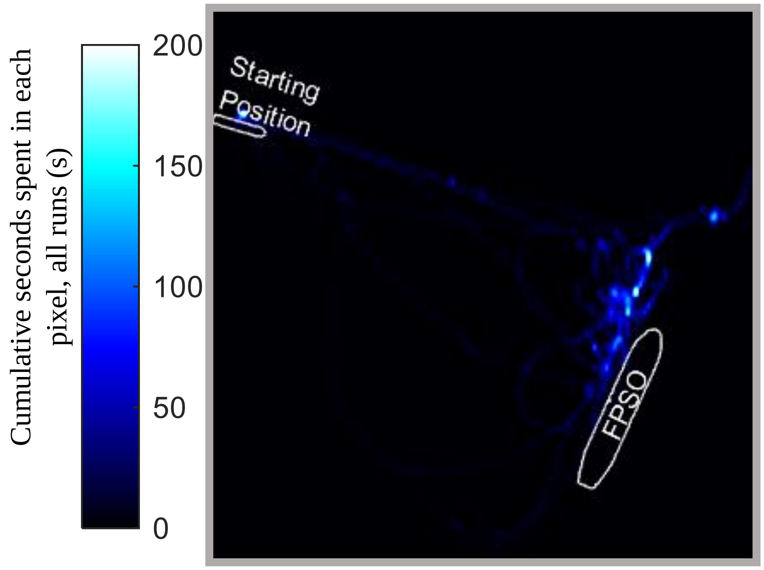


Figure 45: Sample of Heat Map showing starting position. Units are cumulative seconds in pixel, all vessels.

The first comparison examines the DSS group in comparison with the No DSS group and Seafarers in mild ice conditions. It was observed in Figure 46 that the strategies of the three groups do not differ greatly. In Figure 46, increased vessel concentration is seen upstream of the FPSO in the DSS group, as was suggested by the DSS in several scenarios. This is most evident in the DSS group, however some increase in concentration can also be seen in the Seafarers group. The No DSS group is primarily concentrated near the bow of the FPSO; participants also travelled upstream, forward of the FPSO.

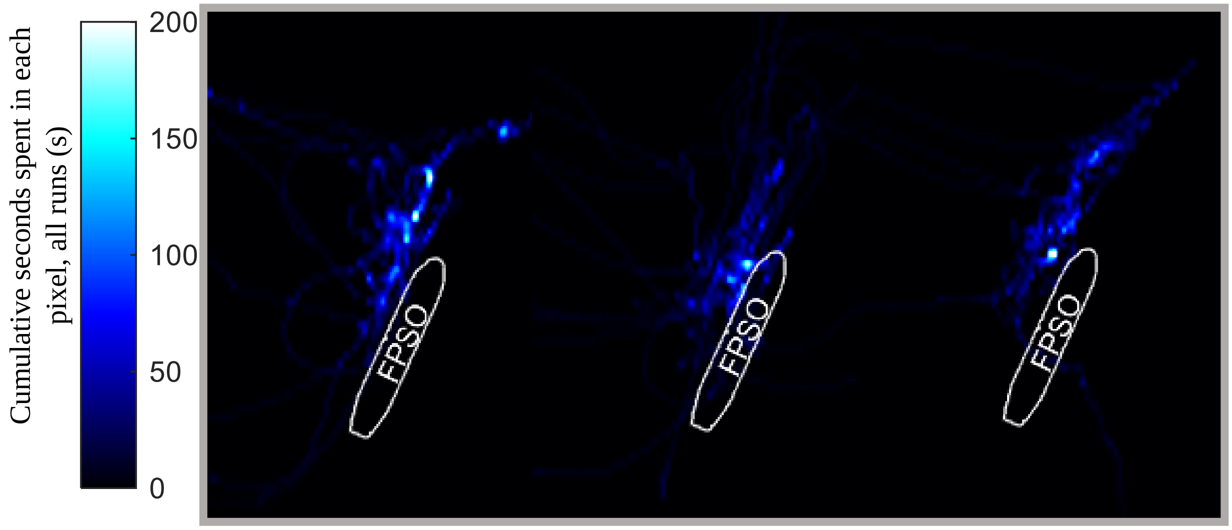


Figure 46: Heat maps for DSS group (left) and No DSS group (middle), and the Seafarer group (right) in mild ice conditions.

Next, the heat maps for the Training I and Training II groups are shown alongside the DSS group in Figure 47. In this, it can be seen that the Training II group has a similar concentration upstream of the FPSO to the DSS group. The Training I group is more spread out and has no clear concentration. It should be noted that the Training I group only had 8 participants, while all other groups presented here had 9. This was due to a saturation of the recruitment pool during Experiment II (Thistle , 2019).

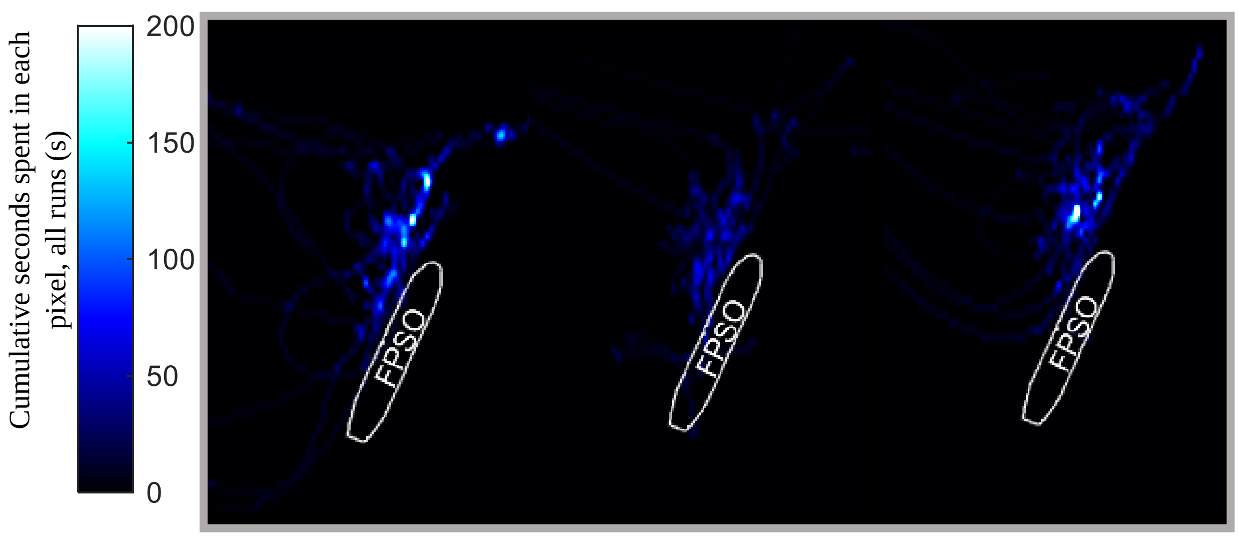


Figure 47: Heat map showing DSS group (left), Training I group (center), and Training II group (right) for mild ice conditions.

Examining the severe ice condition group in Figure 48, it can be seen that there is clustering on the port side of the FPSO in the DSS group and Seafarer group that is not present in the No DSS group. This positioning was considered to be a good strategy as the participants can use the current and their vessel to block ice floes. This strategy was communicated by the DSS and was known by the experienced seafarers, many of whose runs were presented in the DSS to new participants.

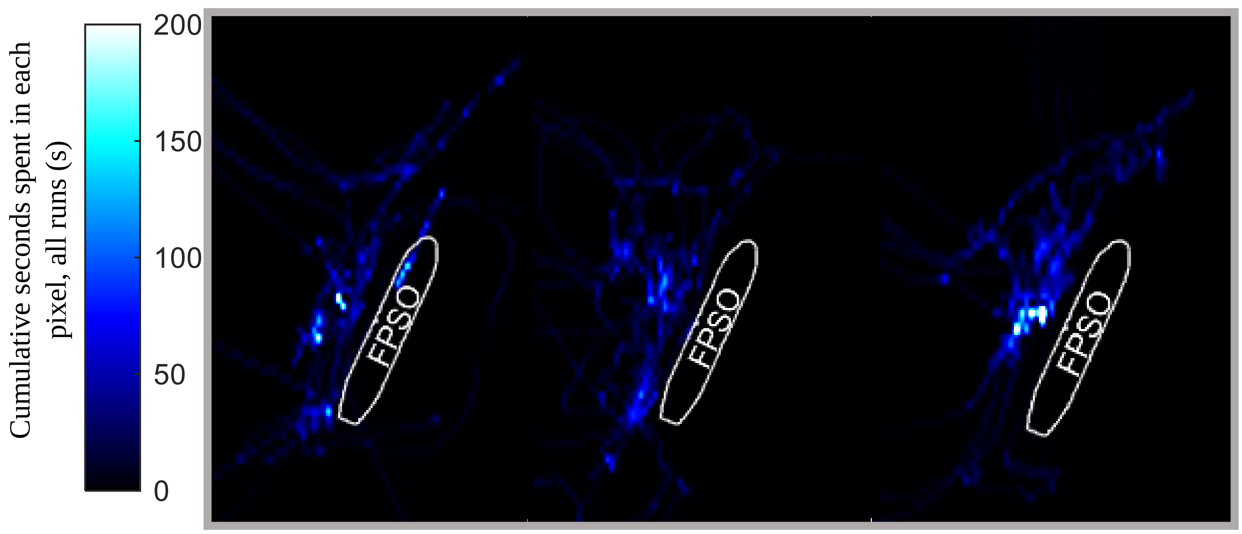


Figure 48: Heat map showing DSS group (left), No DSS group (center), and Seafarers group (right) for severe ice conditions metric.

Examining the high ice concentration group in Figure 49, it can be seen that there is a distinct clustering on the port side of the FPSO near the lifeboat launch zone in all groups. This is most concentrated in the Training I and Training II groups, however it is present in the DSS group as well. This is in contrast to the No DSS group, shown in the center of Figure 48, where no clustering is observed. This result can also be loosely correlated to the performance metrics, from an observational perspective. The groups with more succinct clustering tended to have a higher performance in most metrics.

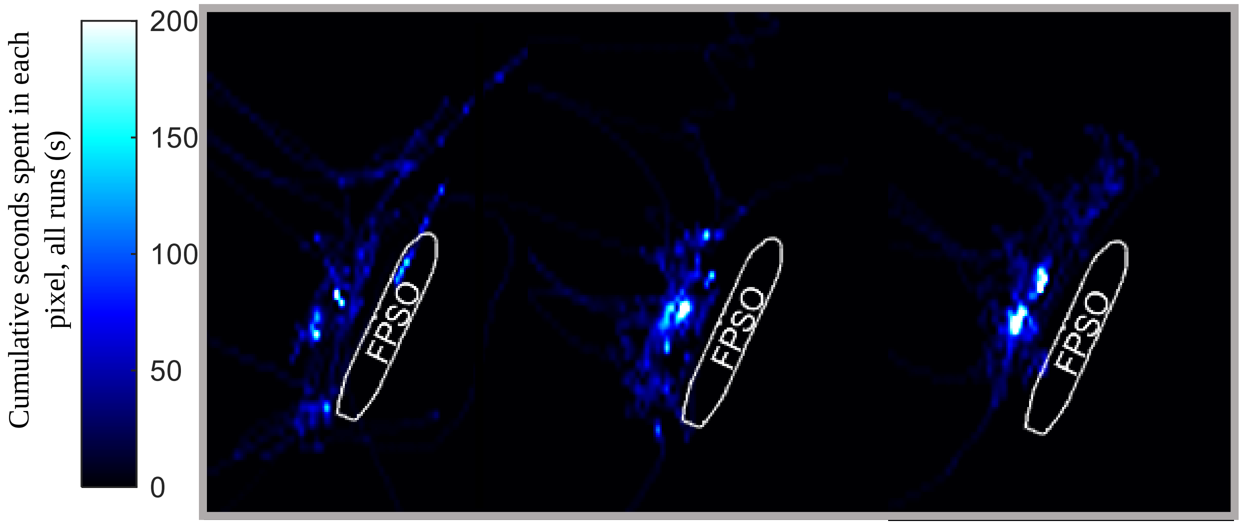


Figure 49: Heat map showing DSS group (left), Training I group (center), and Training II group (right) for the severe ice conditions.

4.4 Comparison of Cadets to Irregular participants

An informal statistical analysis was performed on the group of DSS Cadets from the Nautical Science program to the "Irregular" participants who were recruited from the general student population. It was thought that as many of the Cadets were in the early stages of their programs, with most not having yet been to sea, they would not be significantly different than participants of other technical academic programs. Six additional participants were recruited. As the sample size for this sub-group was small, no power analysis was performed as there was no likelihood of sufficient power to conclusively avoid Type II errors. Thus, these results should be considered an observational justification for including the irregular participants in the DSS group.

In Table 16, it can be seen that in most cases, the irregular participants actually performed better than the cadets in mild ice conditions. The exception to this is in the First Time Clear metric. A *t*-test was also performed to check significance. Again, this must be considered observational as it is not possible to assume normality for a group with so few participants, however it shows that the difference is significant in the case of the Mean Concentration Change, Mean Total Clearing, and Mean Clearing to Distance metrics. It is unlikely that there are any factors that would contribute to the superior performance of the irregular participants over the cadets, therefore it is assumed that this is due to chance that the participants chosen were naturally skilled at the exercise, and had a degree of luck. This indicates that with no significant advantage found in participants from the nautical science program, including participants from the general university population is a valid way to improve statistical power.

Table 17 provides a similar comparison for participants in the severe ice concentration group. In all metrics, with the exception of the Max Time Clear, and Cumulative Time Clear metrics, the cadet sub-group performed better than the irregular sub-group. However, in examining the *t*-test result for this comparison, it can be seen that for all metrics, the *p*-value is significantly greater than 0.05, therefore the difference in performance is not found to be statistically significant. This further validates the decision to include the irregular participants in the sub-group.

Table 16: Comparison of DSS sub-groups for mild ice concentration metric.

Means Cadet to Irregular (4/10th)	**Mean Con Change (X/10th)**	**Max Con Change (X/10th)**	**Mean Total Clearing (km^2)**	**Max Time Clear (s)**	**Cumulative Time Clear (s)**	**Mean Clearing to Distance ratio**	**First Time Clear (s)**
DSS Cadet	1.32	3.70	0.61	245	427	0.40	934
DSS Irregular	2.60	5.92	1.19	537	737	1.09	930
t-test	0.035	0.105	0.038	0.139	0.115	0.016	0.977

Table 17: Comparison of DSS sub-group for mild ice concentration metric.

Means Cadet to Irregular (7/10th)	**Mean Con Change (X/10th)**	**Max Con Change (X/10th)**	**Mean Total Clearing (km^2)**	**Max Time Clear (s)**	**Cumulative Time Clear (s)**	**Mean Clearing to Distance ratio**	**First Time Clear (s)**
DSS Cadet	1.91	5.39	0.87	290	460	0.53	848
DSS Irregular	1.45	4.29	0.67	353	463	0.41	1300
t-test	0.79	0.71	0.78	0.88	0.70	0.94	0.95

4.5 Comparison of Cadet Experience

As the DSS seemed to influence strategy and a tendency of performance improvement was observed by the DSS group using it, it begs the question as to why effect was small and why no significant improvement was observed. Although no significant difference was found in the nautical science cadet sub-group over the Irregular participant subgroup when using the DSS, it was thought that the higher level of cadet experience in the No DSS group had the potential to positively skew the No DSS performance. In the DSS group, all participants who were nautical science cadets were in their third-year or less: only one participant had experience at sea on a cadet term, and this sea term was in the fall season when no ice was present. Of the Irregulars, none had spent time at sea for academic purposes. This is in contrast to the No DSS group in Experiment I, where 7/18 participants were in fourth year and had been to sea for their program. Many of them had experience in ice. The more experienced cadets in the No DSS group are referred to in this section as the No DSS High Exp. The less experienced cadets are referred to as the No DSS Low Exp.

A potential consequence of this is that the performance effects of the DSS are not as dramatic when compared against No DSS participants who may have had experience performing or observing similar tasks. It is also possible that the No DSS group had more ship handling experience, either on their sea-terms, or in simulation courses. This would result in the No DSS control group potentially having a higher baseline.

A breakdown of these sub-groups was performed. The results can be seen in Table 18 and Table 20, with the *t*-tests showing the statistical significance of these comparisons shown in

Table *19* and Table 21. In Table 18, it can be seen that generally, the DSS group performs worse than the No DSS High Exp in most metrics for the mild ice concentration group. Examining the No DSS Low Exp group, it can be seen that the DSS group performs better in all metrics. In the comparison between the DSS and No DSS High Exp group, the results are mixed, with the DSS group performing better in approximately half the metrics. As discussed previously, there is a tendency for the DSS group to perform better than the No DSS group, overall however these results are not statistically significant. Most of the comparisons between the DSS group and the No DSS sub-groups are also not statistically significant (Table 21).

Table 18: Cadet Experience subgroup summary of mean performance metrics for mild ice conditions.

Group	**Mean Conc Change (X/10th)**	**Max Conc Change (X/10th)**	**Mean Total Clearing (km^2)**	**Max Time Clear (s)**	**Cumulative Time Clear (s)**	**Mean Clearing to Distance ratio**	**First Time Clear (s)**
DSS	1.75	4.44	0.80	342	530	0.63	1029
No DSS High Exp	1.30	4.82	0.60	675	918	0.36	590
No DSS Low Exp	0.93	3.87	0.42	174	320	0.17	1240
No DSS Full Group	1.10	4.29	0.50	397	586	0.25	951

Table 19: *t*-tests for comparisons between cadet experience sub-groups, mild ice conditions.

Group	**Mean Conc Change (X/10th)**	**Max Conc Change (X/10th)**	**Mean Total Clearing (km^2)**	**Max Time Clear (s)**	**Cumulative Time Clear (s)**	**Mean Clearing to Distance ratio**	**First Time Clear (s)**
DSS to Low Exp No DSS	0.10	0.53	0.09	0.21	0.15	0.04	0.28
DSS to High Exp No DSS	0.42	0.74	0.41	0.15	0.06	0.26	0.03
DSS to Full group	0.10	0.47	0.54	0.41	0.40	0.26	0.56
High Exp to Low Exp	0.43	0.24	0.43	0.07	0.01	0.06	0.01

In the severe ice concentration category (Table 20), a similar trend can be observed, with the No DSS High Exp sub-group performing better than the No DSS Low Exp subgroup. In several metrics, the DSS group can be seen to perform worse than the No DSS High Exp subgroup, while the trend is reversed for the Low Exp subgroup. Again however, the results are not statistically significant as can be seen by the *p*-values greater than 0.05 in Table 21.

Table 20: Cadet Experience subgroup summary of mean performance metrics for mild ice conditions.

Group	Mean Con Change ($X/10^{th}$)	Max Con Change ($X/10^{th}$)	Mean Total Clearing (km^2)	Max Time Clear (s)	Cumulative Time Clear (s)	Mean Clearing to Distance ratio	First Time Clear (s)
DSS	1.76	5.02	0.81	311	461	0.49	999
No DSS High Exp	2.19	5.42	1.00	420	573	0.61	1203
No DSS Low Exp	1.70	5.29	0.78	232	325	0.41	1150
No DSS Full Group	1.87	5.33	0.86	294	408	0.48	1168

Table 21: *t*-tests for comparisons between cadet experience sub-groups, severe ice conditions.

Group	Mean Con Change ($X/10^{th}$)	Max Con Change ($X/10^{th}$)	Mean Total Clearing (km^2)	Max Time Clear (s)	Cumulative Time Clear (s)	Mean Clearing to Distance ratio	First Time Clear (s)
DSS to Low Exp No DSS	0.61	0.90	0.61	0.42	0.17	0.38	0.20
DSS to High Exp No DSS	0.47	0.97	0.46	0.50	0.61	0.62	0.28
DSS to Full group	0.91	0.93	0.91	0.97	0.71	0.67	0.15
High Exp to Low Exp	0.45	0.89	0.45	0.34	0.28	0.16	0.87

Despite the lack of significance, it can be observed that there is a tendency for the more experienced subgroup of Cadets used for the No DSS control group to perform the ice management tasks more effectively than the average of this group. All participants in the DSS group are considered to be Low

Experienced, meaning few of them had been to sea, and none had the same exposure to ship handling and ice management as those in the High Experienced No DSS group. There is a chance that this skewed the results, causing a comparatively lower performance in the DSS group than if more of the participants had been in the later years of a nautical science program. It is expected that the DSS group would have performed better if they'd had experience levels distributed similarly to the No DSS group. Due to a saturation of the recruitment pool, no highly experienced participants were able to be recruited from this group.

4.6 Qualitative Results

Following each participant's run, an exit interview was performed. The results of this can be seen in the appendix, section 7.2. The purpose of this was to gauge which components of the DSS were found to be the most useful the participants, and which components were not used. A summary of this can be found in Table 22. Referring back to Figure 8, the Video and Picture in Table 22 refer to the Video and Picture on the DSS display. The Text refer to the text based instructions. Numbers refers to the numerically based "Suggested Solution" component of the DSS. From this, it can be seen that participants found the video to be the most useful component of the DSS, with 14/18 reporting that they found it to be the most useful. The text and numbers were found to be the least useful. In some cases, no data was collected or no opinion was given. This is presented as "N/A", or Not Applicable.

Additionally, a self-rating of performance was done by the participants, as well as a rating of the usefulness of the DSS. Both of these ratings were on a scale of 1-5. The mean self-performance rating was 3.33/5. The mean rating of the DSS usefulness was 3.97/5. Details of this can be seen in Table 22.

Finally, the number of times the participants requested assistance was also recorded, and is also presented in Table 22. From this, it can be seen that there is a high rate of variability in between participants, with some only requesting assistance once, and one requesting assistance up to 19 times. The average number of assistance requests was 5.

Table 22: Summary of self-rating, assistance requests, and most/least useful components of the DSS

Participant	Self Rating	DSS Rating	Assistance Requests	Most Useful	Least used
1	2	4	3	Video	Text
2	2	3	2	Video	Numbers
3	4	5	1	Video	Text
4	3	4	1	Picture + Text	Video
5	3.4	5	5	Video + Picture	Text
6	3	4.5	4	Text	Video
7	3	3	2	Text + Video	Numbers
8	3	4	29	Video	Text
9	5	4	1	Video	Text + Numbers
10	3	4	4	Video + Text	Pictures
11	5	5	1	Videos	Text + Numbers
12	4	4.5	6	Numbers	Picture
13	1.5	2.5	5	Video + Numbers + Pictures	Text
14	4	5	1	Video	N/A
15	3	3	1	Picture	N/A
16	4	5	6	Video	Numbers
17	5	3	9	Video	Numbers
18	2	3	3	Video	N/A
Mean	3.33	3.97	4.70		

Also, the participants were asked what their planned strategy was, and how it panned out. This was useful to see whether the DSS provided guidance that was used by the participants or not. All participants stated that the DSS was useful to them for choosing a strategy. Data was also gathered showing how many times each participant used the DSS in a run, at which time interval, and what strategy was recommended.

In addition to this, participants performed an experience questionnaire. An example of this can be found in Table 24 and Table 25 of the appendix. The purpose of this was to gauge whether the participants had any experience handling vessels, experience at sea, or experience in simulators. As all participants were recruited from an "inexperienced" pool, this experience was naturally limited. One participant had performed a co-op term on an offshore construction vessel, but had not encountered sea ice. Several others had worked on small vessels as summer jobs, or with family. The rest had no marine experience outside of class.

Discussion and Future Work

4.7 Design Power and Recruitment Difficulties

One of the weaknesses of this experiment was the lack of sufficient design power among most of the performance metrics. The consequence of this is that although observational conclusions can be made based on the numerical results, it is possible that a Type II error has been made and that the null hypothesis that no significant difference exists between the DSS and No DSS groups has been falsely accepted. The original a-priori calculations for design power estimated that the effect size between the DSS and No DSS groups would be the same as the effect size in Experiment II's Training I and No Training groups (Thistle & Veitch, 2019). In that study, the effect size for the a-priori power calculations was estimated as being the same as the effect size between the Cadet and Seafarer groups in Experiment I (Veitch et al., 2018). In both cases, the estimated effect was found to materialize in practice, leading to sufficient design power in the experiment. Due to a lack of additional literature on the effects of using a DSS for offshore ice management, the same effect was estimated for this Experiment III. Upon completing the experiment, the effect was found to be less than had been predicted, and consequently there is insufficient design power to be able to conclusively state no Type II errors exist, as was shown in Table 3.

Assuming the actual effect size does not vary considerably if more participants were recruited, a new power analysis was completed in an attempt to determine what number of participants would be required to avoid the risk of a Type II error. The calculations were once again performed using the G*Power software. First, the calculation was performed assuming an effect size of 0.84. This was the largest effect found between the DSS and No DSS groups for the Mean Clearing to Distance Ratio metric. In Figure 50, it can be seen that a sample size of 47 would be required to achieve a power of 0.8. In order to maintain symmetry between groups, an even sample size would be required, meaning 48 participants, or 24 per group. This would require an additional recruitment of 12 participants, or 6 per group.

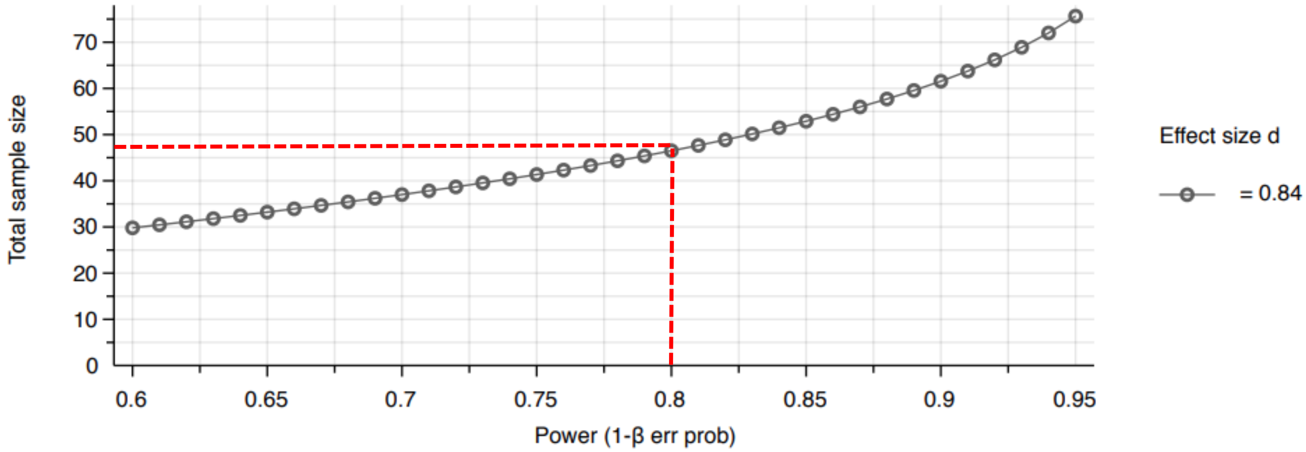


Figure 50: Power analysis for hypothetical future recruitment in Clearing to Distance Ratio metric showing a sample size of 48 participants required to achieve a design power of 0.80.

As increasing the recruitment to 24 per group would only result in sufficient design power for the one metric that showed the largest effect from the DSS, an analysis was also performed on the smallest effect to see how many participants would need to be recruited for all metrics to have 80% design power. The smallest effect was found in the Max Time Clear metric, shown in

Table *2*. This also resulted in the lowest design power, shown in Table 3. The effect was found to be 0.16 using Cohen's *d* method. Assuming that this effect would remain the same should the experiment continue, G*Power was used to calculate the required sample size to achieve 80% power in this metric. From Figure 51, it can be seen that a sample size of 1230 participants would be required in total, or 615 per group. This is an unrealistic and impractical number of participants due to the time and resources which would be required.

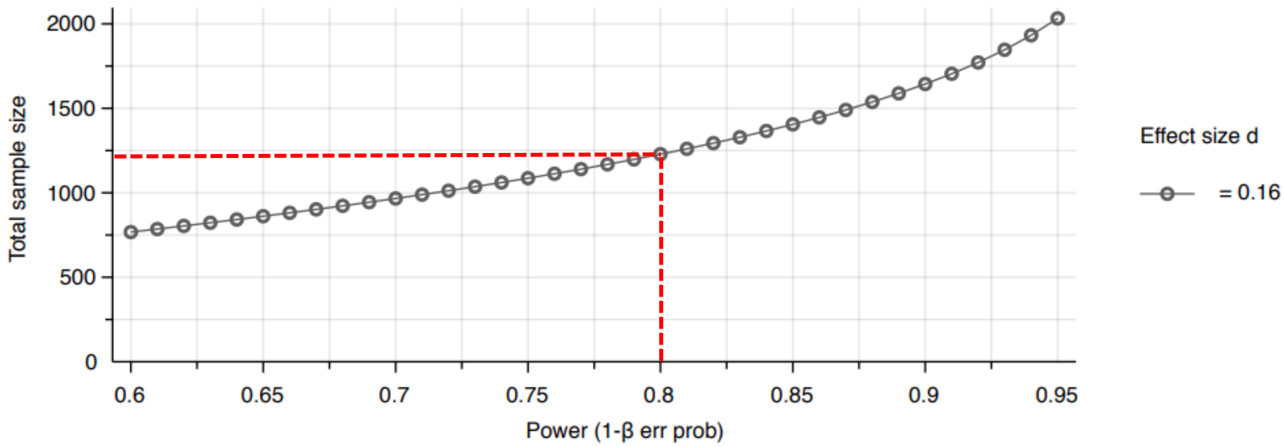


Figure 51: Power analysis for hypothetical future recruitment in Max Time Clear metric showing a sample size of 1230 participants required to achieve a power of 0.80.

The control group would also have had to be expanded, and since the control group dataset was collected in Experiment I, increasing these participants with a new researcher could result in an additional confounding variable. The original pool of participants for the DSS group who were exclusively enrolled in a Nautical Science program was exhausted at 12 participants despite repeated recruitment efforts. This was six short of the desired 18 expected to be required for sufficient design power. To continue recruitment, the pool was widened to include any post-secondary student. The decision to expand the pool was made in December of 2021. A resurgence in the COVID-19 pandemic necessitated a four-week delay in resuming recruitment. Further hardware issues with the experimental simulator resulted in an additional six-week delay in recruitment, as a critical component had to be repaired. It was still possible to recruit the six additional participants in short order once the component was reinstalled, so that the intended sample size of 18 per group could be reached. After analysis of the results, it became apparent that although the increased recruitment improved the design power, the desired level of power was still not reached.

4.8 Quality of Result Metrics

Seven different numerical metrics were calculated in an attempt to gain a clear picture of what "good ice management" performance means. Some of these metrics seem to provide a more clear result, and some were perhaps not good indicators of performance, since differences between groups cannot be observed. The maximum ice concentration metric is perhaps one of these. This metric provides a snapshot of the maximum ice cleared, but does not provide clear indication of the overall quality of a run. For example, a participant could theoretically clear the majority of the ice management zone on their first approach, and fail to maintain any ice-free status at all for the rest of the run. This would provide them with a high maximum ice clearing value, but all other metrics could be poor. One way of determining which metrics were effective and which were not could be to examine the effect sizes and eliminate the metrics that stand out from others as having a low effect. For example, when comparing the DSS to No DSS groups, the Max Concentration Change metric mentioned above, and the Max Time Clear metrics all had what would be considered negligible effects, below 0.20 (Cohen J., 1992). Considering that the other metrics showed at least a small effect, it may be that metrics with negligible effects are simply ineffective and should not be considered in future analyses.

4.9 Participant Strategy

Through observations of the replay videos, the heat maps, and the exit interviews, it is clear that all participant cohorts, with the exception of the No DSS group had a tendency to position themselves in their own-ship up-current from the lifeboat launch zone and use their vessel to block ice from entering the zone. The No DSS group, being the least experienced and not being provided with training or a DSS had a less coherent and concise strategy, and tended to be quite varied in their approaches. They also tended to have the lowest results in most metrics from an observational standpoint. Although the DSS group appeared to use strategies more closely resembling those of the Training I and II groups, and the Seafarers group, this did not translate to significant performance improvements. If the DSS group demonstrated similar tactics to those employed by experienced seafarers and inexperienced participants who had been provided with training, it begs the question as to why the DSS group did not have significant performance improvements.

A closer examination of the replay videos, where a bird's eye view of the 30 minute scenario is sped up by 30, may provide a partial explanation. In every recommendation of the DSS, it is recommended that the participants use the Leeway technique, holding their vessel in place using side thrusters on a heading perpendicular to the FPSO to block ice floes as they flow south with the current. This is a technique that is commonly used in the Seafarers group, and in the Training I and Training II groups. Generally, participants who use this technique were found to have an above average performance. When the DSS group tried to follow these expert strategies, they were slower in their responses and appeared to be uncomfortable getting as close to the FPSO as the aforementioned groups. Many also stated that they were nervous about travelling too fast through the ice (they were instructed to keep speed below approximately 3kts while in ice), and they did not know how close they could safely place their vessel in relation to the FPSO. This resulted in many of the participants maintaining excessive distance between their own-ship and the FPSO, which allowed ice floes to pass between them and the FPSO and into the lifeboat launch zone. Additionally, the more cautious speeds which were often significantly below 3kt, resulted in a delayed response time and resulted in a lower score in many of the performance metrics. An example of this was shown in Figures 4 and 11.

At the beginning of the run and before entering the pack ice, the experienced seafarers were much more likely to have developed a plan or strategy, and had the capability and experience to implement a secondary approach if they found the initial tactics to be ineffective (Smith et al., 2020). The Training I and Training II groups were told the most effective ice management techniques as a part of the training curriculum developed by Thistle (Thistle , 2019). This allowed many of them to formulate and implement a plan based on expert techniques (Thistle & Veitch, 2019). The participants in the DSS group were also able to formulate a plan after seeing the advice of the DSS, however many stated that they had trouble implementing this plan in practice as it was their first time using the simulator in ice, with the exception of a 5 minute habituation scenario relating to ice.

4.10 Outlier Participant

One of the participants had significant difficulties in their run and failed to clear any ice from either zone. This resulted in a score that was equivalent to no ice management being completed at all. The result of this is that in some cases, the assumptions required for the design of experiments was at times violated by this one point. According to Montgomery, a small to moderate violation of assumptions is not a major concern, however a failure of the independence assumption and strong assumptions of non-normality should not be ignored (Montgomery, 2013, p. 42). As there was no clear violation of the assumption of normality in the normal plot of residuals for all metrics, this outlier was not thought to be of concern.

The outlier was the only one of the participants who failed to clear any ice from the ice management zones throughout the scenario, and thus had a score considerably lower than the next lowest participant. In their exit interview, the participant stated that they forgot how to use the side thrusters, and were thus too afraid to use them. This resulted in them having considerable trouble executing the tactics suggested by the DSS. For example, they were attempting to clear ice using leeway, and make slow speed manoeuvers in ice to get themselves near the ice management zone. The issue was that they descended too far down-current at the onset of the run, and so were below the ice management zone. Without using the thrusters, they were unable to turn their vessel since the rudder requires thrust from the main propellers, and they were concerned about using the propellers for fear of reaching too great a speed in ice. This resulted in them spending the majority of the scenario down-current from the target area, and being unable to perform any effective ice management. An example of their performance can be seen in the Change in Ice Concentration plots, shown in Figure 52 and Figure 53. From this, it can be seen that no effective ice management was completed.

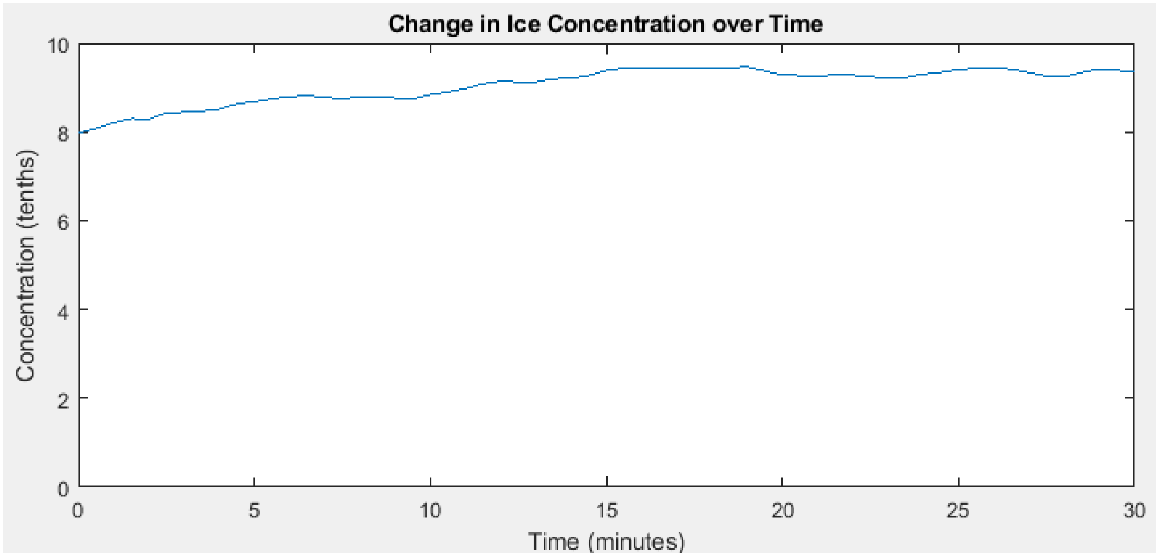


Figure 52: Change in ice concentration in ice management zone for outlier participant showing that no reduction was made.

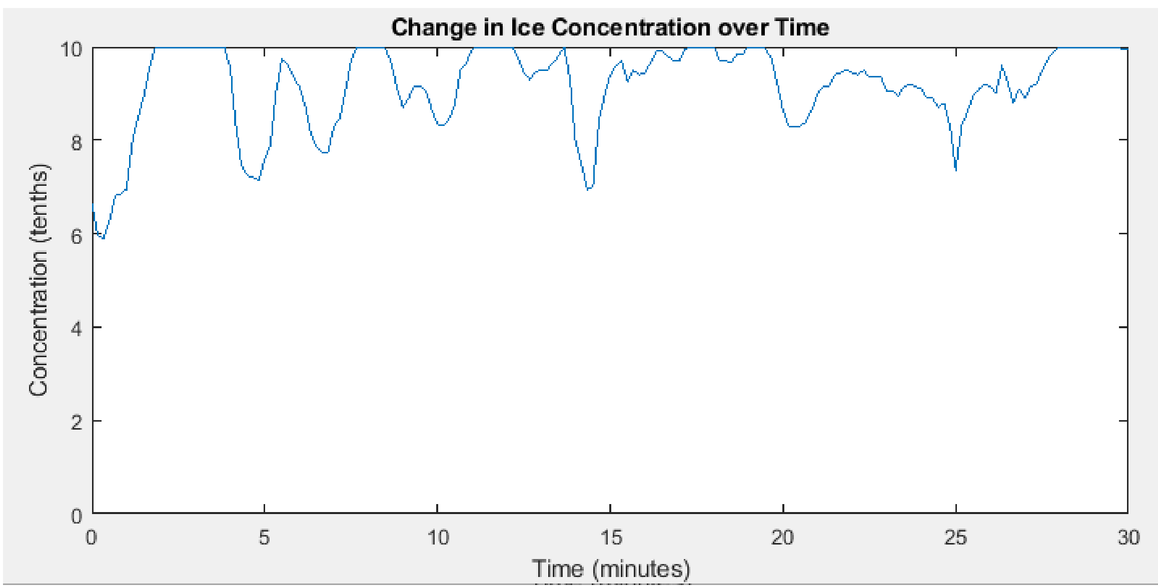


Figure 53: Change in ice concentration in lifeboat launch zone for outlier participant showing that no reduction was made by the participant's actions.

The DSS provides strategic advice and suggests using thrusters at times, however the participant would have had to read all sections of the DSS to be aware of this. In their exit interview, they stated that they did not find the text useful, so they likely did not see this, focusing instead primarily on the video. They also stated that they would prefer more specific instructions on using the controls combined with the strategic advice.

While consideration was given as to whether this outlier should be removed from the analysis entirely, ultimately it was decided that this would not be appropriate. Reasons for removing the outlier were that as all other participants utilized all controls effectively, it was perhaps not realistic to assume that a navigation officer, even one who is inexperienced, would fail to utilize a critical tool for ship-handling at low speeds. The participant was one of the irregular participants. This caused some concern that the decision to widen the recruitment pool could negatively affect the results. Additionally, removing the outlier would have strengthened the validity of the assumptions and resulted in significance of the ice concentration factor for some metrics.

There were several reasons for not removing the outlier. As the experiment was focused on providing inexperienced participants with a DSS, there was no reason why this inexperienced participant should be ineligible. Additionally, other participants from the irregular pool outside of the nautical science program performed very well, often better than students who were enrolled in the nautical science program. Thus it was decided that the poor performance of one irregular participant, the outlier, was not sufficient grounds to question the increasing of the recruitment pool. Also, as the violation of the assumptions was considered to be mild and could effectively be ignored, the benefits of removing this outlier were hard to justify. The balance of this reasoning was that the outlier was not removed from the analysis.

4.11 Equipment Design

Results of the exit interviews for the DSS participants provided important feedback as to how the design of the DSS affected performance. Observations from the interviews with the DSS participants indicated that the video replays in the GUI was preferred over the text and numerical based information. A breakdown of the most and least used components is shown in Table 22. Participants stated that they generally found the text-based instructions too distracting, and that the overall GUI was cluttered. This led to the participants gravitating towards the looping video in the center of the display as the information being presented was concise and clear. Participants who did read the text-based instructions in conjunction with the video or picture did find the information helpful as each bullet point was customized to the specific scenario being presented. Future work should take care to incorporate the best of human-machine interface design (HMI), and avoid cluttering the GUI with too much data (Lee et al., 2017). This could also include consistent design so that the data presented in the DSS matched the data presented in the other simulator systems (Nordby et al., 2019).

Improving the participant's understanding of the DSS design and functionality could also be beneficial. This would allow participants to have a better sense of the strengths and limitations of the system. Bainbridge's Ironies of Automation supports this idea by discussing how lack of understanding of automated systems in all industries can threaten the safe operations of said systems (Bainbridge, 1983). This is also supported in work that examined factors leading to misunderstandings of course predictor software in advanced bridges, where a lack of understanding behind data sources and uses could lead to marine incidents by pilots (Westin et al., 2019). A study by Nilsson et al. (2009) examined the effects of advanced technology on the safety of marine operations. It was found that experienced navigation officers performed better with conventional ship's bridges than with technically advanced bridges when performing simulated fairway navigation scenarios. This trend was reversed with inexperienced officers (Nilsson et al., 2009). If the DSS equipped bridge is considered technologically advanced, a future study could gauge the effects of a DSS on experienced participants as well as less experienced participants in an attempt to determine whether the DSS effects found here hold true for more experienced participants. This would be performed after setting a new benchmark with an improved DSS.

4.12 Future Improvements

By examining the common recommendations in the exit interviews and the findings in literature, improvements in the DSS design can be identified. A list of the recommendations by participants can be found in Table 23. Future improvements to the DSS would incorporate a more user-friendly method for requesting assistance, such as a physical button instead of a trackpad. This would be consistent with best practices of designing with regards to human-factors, since haptic feedback can be beneficial to users (Lee et al., 2017). This would reduce the need for a user to remove their attention from the task at hand to search for a virtual button on the monitor, which must be pressed using a trackpad. The use of physical buttons over virtual ones is also supported and encouraged by the findings form the National Transportation Safety Board of the marine accident of the U.S.S. John S McCain. Some of the findings of the report stated that an overreliance on digital throttle controls on a touch screen likely led to a loss of situational awareness by the navigational crew, contributing to a collision (National Transportation Safety Board, 2017). This incident led to an announcement by the US Navy that it planned to remove all touch screen throttle and helm controls across the entire fleet, replacing them with more traditional physical controls (Mallam et al., 2020).

Table 23: Participant recommendations for improvement

Participant	Future Improvement Recommendations
1	More specific instructions
2	Being able to replay/indefinitely loop video
3	More colours in instructions and highlighting key words
4	Radar + AIS overlay
5	Overlay past track
6	Larger video
7	Continuously looping video, more specific details
8	Automatic assistance
9	Step-by-step pictures and less text
10	Step-by-step diagrams instead of a single picture to complement instructions
11	Slow video
12	Slow video
13	Add verbal instructions and give more specific controls advice
14	Remove bad scenarios
15	N/A
16	Allow cycling through scenario recommendations
17	Allow cycling through scenario recommendations
18	Increase pixel count of ship's path, allow slowing of video

Several participants also stated that it would be helpful to have the ability to play, pause, and control the speed of the replay video. Currently, the video plays at 300 times the original speed, in a six-second loop. Several participants stated that this was too fast, and that having the ability to pause or slow the playback would be helpful. In keeping with the best practices discussed above, a simple physical knob control could be provided so that easy playback speed adjustments could be performed. A physical play/pause button could also be included. This would provide all participants with the opportunity to partially customize their user experience.

There are risks associated with designing a bridge in such a way that many of the components do not match each other in style or feel (Nordby et al., 2019). This is referred to as a multivendor bridge system, and risks creating inconsistencies that can alter the situational awareness of users (Nordby et al., 2019).

Future work would also seek to reduce text on the DSS or replace it with pictograms, or voice commands. This would reduce the risk of participants removing their attention from the task at hand to read text on the screen, which was a common complaint by users in their exit interviews.

Finally, improvements should be made to the CBR algorithm in the DSS. If a participant's approach is largely different from an approach in the base, there could be situations in which the algorithm suggests an approach that is unrealistic or difficult to achieve given the current position of the participant. This has the potential to be largely unhelpful, or even confusing for the participant, and is a failure of automation reliability, where the automated system leads to distrust by the user due to its perceived poor performance (Lee et al., 2017). In at least one case, a user was provided with a suggestion like this, and the decided to no longer use the DSS for the remainder of the task.

5 Conclusions

This research ultimately provided evidence that although the quantitative effects are limited, the DSS can influence the behaviour and strategy of participants. It also demonstrated that the experimental method is effective for evaluating a DSS in a simulator environment. Finally, interviews with participants in conjunction with literature on designing with human factors in mind provided a good indication of what improvements could be implemented in future versions of the DSS.

The first hypothesis that the DSS would result in performance improvements was rejected since no statistically significant improvement was found. Thus, the null hypothesis that no significant difference exists must be accepted. However, it must also be noted that the assumptions for the effect size of the DSS performance were incorrect, resulting in a lower design power than expected. This renders a conclusive argument that the results are statistically insignificant untenable, since the risk of a Type II error that the null hypothesis is falsely accepted exists. The second hypothesis was that ice concentration would not result in any significant differences in performance. This was accepted in the case of all metrics examined. The implication of this is that ice concentration is not a confounding factor in the results.

Although it did not deliver statistically significant performance improvements, the DSS demonstrated that it is effective at influencing group behaviour and delivering expert strategic advice to inexperienced participants, as shown by the heat maps, observational improvements in performance, and the exit interviews. Additionally, although the DSS alone may be insufficient to significantly improve the performance of novice users, the effects of increased training to improve ship-handling skills combined with a DSS for strategic emulation may be worth investigating. With improvements, the DSS has the potential to assist in ice management training, and could eventually be used for strategic assistance on board a vessel. The content of the exit interviews in conjunction with literature on human factors design can be used to improve future DSS projects.

6 References

Bainbridge, L. (1983). Ironies of automation. *Automatica, 19*(6), 775-779. https://doi.org/10.1016/S1474-6670(17)62897-0

Canadian Coast Guard College. (2018). Ice Management Simulator Lesson Plans and Scenarios .

Cohen, B. H., & Lea, B. R. (2004). *Essentials of Statistics for the Social and Behavioural Sciences* . Hoboken: Wiley .

Cohen, J. (1969). *Statistical Power Analysis for the Behavioural Sciences* . New York: Academic Press.

Cohen, J. (1992). Statistical Power Analysis. *Journal of Clinical Psychiatry, 51*, 61-69.

Corder, G. W., & Foreman, D. I. (2014). *Nonparametric Statistics: a Step-by-Step Approach* . Hobokin: Wiley.

Dunderdale, P., & Wright, B. (2005). *Pack ice management on the southern Grand Banks offshore Newfoundland, Canada.* Noble Denton Canada Ltd.

International Maritime Organization. (2010). Guidelines for Ships Operating in Polar Waters.

International Maritime Organization. (2017). Model Course 7.11: Basic Trainig for Ships Operating in Polar Waters.

International Maritime Organization. (2021). *Outfome of the Regulatory Scoping Excercise for the Ise of Maritime Autonomous Surface Ships (MASS).* London: International Maritime Organization. Retrieved from https://wwwcdn.imo.org/localresources/en/MediaCentre/HotTopics/Documents/MSC.1-Circ.1638%20-%20Outcome%20Of%20The%20Regulatory%20Scoping%20ExerciseFor%20The%20Use%20Of%20Maritime%20Autonomous%20Surface%20Ships...%20(Secretariat).pdf

International Organization for Standardization. (2018). ISO 35104: Petroleum and natural gas industries - Arctic operations - Ice management.

Kolodner, J. L. (1992). An Introduction to Case-Based Reasoning. *Artificial Intelligence Review, 6*, 3-34.

Lee, D. J., Wickens, C. D., Liu, Y., & Boyle, N. L. (2017). *Designing for People: an Introduction to Human Factors Engineering.* Charleston, South Carolina: CreateSpace.

Mallam, S. C., Nazir, S., & Sharma, A. (2020). The human element in future Maritime Operations - perceived impact of autonomous shipping. *Ergonomics, 63*(3), 334-345. https://doi.org/10.1080/00140139.2019.1659995

Mallam, S. C., Nordby, K., Johnsen, S. O., & Bjørneseth, F. B. (2020). The digitalization of navigation: Examining the accident and aftermath of US Navy destroyer John S. McCain.

Microsoft Corporation . (n.d.). *Create a box and whisker chart*. Retrieved from Microsoft Support: https://support.microsoft.com/en-us/office/create-a-box-and-whisker-chart-62f4219f-db4b-4754-aca8-4743f6190f0d

Montgomery, D. C. (2013). *Design and Analysis of Experiments.* Hoboken: John Wiley & Sons Inc.

National Transportation Safety Board. (2017). *Collision between US Navy Destroyer John S McCain and Tanker Alnic MC.* Washington: National Transportation Safety Board.

Nilsson, R., Gärling, T., & Lützhöft, M. (2009). An experimental simulation study of advanced decision support syste for ship navigation. *Transportation Research Part F, 12*, 188-197. https://doi.org/10.1016/j.trf.2008.12.005

Nordby, K., Mallam, S. C., & Lützhöft, M. (2019). Open user interface architecture for digital multivendor ship bridge systems. *WMU Journal of Maritime Affairs*, *18*, 297-318. https://doi.org/10.1007/s13437-019-00168-w

NVIDIA Corporation. (2008). NVIDIA PhysX System Software Version 2.8.1. Santa Clara, California, U.S.A. .

Pazouki, K., Forbes, N., Norman, R. A., & Woodward, M. D. (2018). Investigation on the impact of human-automation interactin in maritime operations. *Ocean Engineering*. DOI:10.1016/j.oceaneng.2018.01.103

Pietrzykowski, Z., Piotr, W., & Borkowski, P. (2017). Decision Support in Collision Situations at Sea. *The Journal of Navigation*, 447-464. DOI:10.1017/S0373463316000746

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin.

Sharma, A., Kim, T.-e., & Nazir, S. (2019). Catching up with time? Examining the STCW competence framework for autonomous shipping. *Ergoship.* Haugesund.

Simões Ré, A., & Veitch, B. (2007). Lifeboat operational performance in cold environments. *Design and Construction of Vessels Operating in Low Temperature Environments.*

Sivia, D. S., & Skilling, J. (2006). *Data Analysis: A Bayesian Tutorial.* Oxford: Oxford University Press.

Skjong, S., Kyllingstad, L. T., Reite, K.-J., Haugen, J., Ladstein, J., & Aarsaether, K. G. (2019). Genetic on-board decision support system framework for marine operations. *Conference on Ocean, Offshore, and Arctic Engineering.* Glasgow: ASME.

Smith, J., Yazdanpanah, F., Thistle, R., Musharraf, M., & Veitch, B. (2020). Capturing Expert Knowledge to Inform Decision Support Technology for Marine Operations. *Journal of Marine Science and Engineering, 8*(9), 689. https://doi.org/10.3390/jmse8090689

Thistle , R. (2019). An Evaluation of the Effects of Simulator Training on Ice. *Master's Thesis*, *Memorial University of Newfoundland*. St. John's, NL, Canada.

Thistle, R., & Veitch, B. (2019). An Evidence Based Method of Training to Targeted Levels of Performance. *SNAME Maritime Convention.* St. John's, NL.

Tsou, M.-C., Kao, S.-L., & Su, C.-M. (2010). Decision Support from Genetic Alrogirhms for Ship Planning and Alerts. *The Journal of Navigation*, 167-182. DOI:10.1017/S037346330999021X

Veitch, E., Molyneux, D., Smith, J., & Veitch, B. (2018). Investigating the Influence of Bridge Officer Experience on Ice Management Effectiveness using a Marine Simulator Experiment. *Journal of Offshore Mechanics and Arctic Engineering*, *141*(4). doi:10.1115/1.4041761

Westin, C. A., Lundin Palmerius, K. E., Johansson, J., & Lundberg, J. (2019). Concept of Reskilling for Automation Collaboration in Maritime Piloting. *International Federation of Automatic Control*. https://doi.org/10.1016/j.ifacol.2019.12.090

Wróbel, K. (2021). Searching for the origins of the myth: 80% human error impact on maritime safety. *Reliability Engineering and System Safety*. doi:https://doi.org/10.1016/j.ress.2021.107942 https://doi.org/10.1016/j.ress.2021.107942

Yazdanpanah, F. (2021). Designing a Cased Based Reasoning Decision Support System for Ice Management Operations Using Expert Knowledge. *Masters Thesis*. Memorial University of Newfoundland.

7 Appendices

7.1 Participant Experience

Table 24: Experience questionnaire part 1

	Concentration	Year of birth	Gender	Year of Program	Are you enrolled in a Nautical Science program?	If no, what program are you enrolled in?	Months spent at sea	Vessel Experience (type)	Experience operating in sea ice (y/N)
1	7	2003	Male	1	Y		0	None	N
2	4	2000	Male	1	Y		0		N
3	7	1980	Female	2	Y		0	0	N
4	4	2002	Male	2	Y		2	Ferry/Coastal	N
5	7	1996	Female	1	Y		0	0	N
6	4	2002	Female	1	Y		55	Coastal/Fishing	N
7	7	2003	Male	1	Y		0	0	N
8	4	2002	Male	2	Y		0	0	N
9	7	1987	Male	3	Y		2.5	OSV/AHTS	N
10	4	2003	Female	1	Y		0	0	N
11	7	1992	Male	2	Y		0	0	N
12	4	2003	Female	1	Y		0	0	N
13	7	1992	Female	3	N	M Eng ONAE	0	0	N
14	4	1997	Male	1	N	M Eng ONAE	0	0	N
15	7	1995	Female	>4th	N	Medicine	0	Small pleasure craft	N
16	4	1993	Male	1st	N	PhD ONAE	0	0	N
17	4	1996	Male	2	N	M Eng ONAE	0	0	N
18	7	1993	Male	3	N	Electrical Eng	1	Ferry/Coastal	N

Table 25: Experience questionnaire part 2

	Participant Code	Concentration	Operations performed in ice	Location of Experience	Years since last operated in sea ice	Years spent in presence of ice	Shore based training for ice operations	Marine simulator experience?	Note
1	**P75**	**7**	**N/A**	**N/A**	**N/A**	**0**	**None**	**None**	
2	**T35**	**4**	**N/A**	**N/A**	**N/A**	**0**	**None**	**None**	
3	**T70**	**7**	**N/A**	**N/A**	**N/A**	**0**	**None**	**Yes**	**training for open water navigation and for a research study**
4	**W25**	**4**	**N/A**	**N/A**	**N/A**	**0**	**None**	**None**	**2nd mate on small tour boat**
5	**C40**	**7**	**N/A**	**N/A**	**N/A**	**0**	**None**	**None**	**Worked on cruise ships in entertainment**
6	**V72**	**4**	**N/A**	**N/A**	**N/A**	**0**	**None**	**None**	**worked on family fishing boats and tour boat**
7	**D84**	**7**	**N/A**	**N/A**	**N/A**	**0**	**None**	**None**	
8	**N17**	**4**	**N/A**	**N/A**	**N/A**	**0**	**None**	**Yes**	**training for open water navigation and for a research study**
9	**B61**	**7**	**N/A**	**N/A**	**N/A**	**0**	**None**	**Yes**	**training for open water navigation**
10	**S63**	**4**	**N/A**	**N/A**	**N/A**	**0**	**None**	**None**	
11	**D86**	**7**	**N/A**	**N/A**	**N/A**	**0**	**None**	**Yes**	**training for navigatio**

									n in open water
12	**L59**	**4**	**N/A**	**N/A**	**N/A**	**0**	**None**	**None**	
13	**A92**	**7**	**N/A**	**N/A**	**N/A**	**0**	**None**	**None**	
14	**H29**	**4**	**N/A**	**N/A**	**N/A**	**0**	**None**	**Yes**	**research study**
15	**E90**	**7**	**N/A**	**N/A**	**N/A**	**0**	**None**	**None**	
16	**F41**	**4**	**N/A**	**N/A**	**N/A**	**0**	**None**	**None**	
17	**R12**	**4**	**N/A**	**N/A**	**N/A**	**0**	**None**	**None**	
18	**O52**	**7**	**N/A**	**N/A**	**N/A**	**0**	**None**	**None**	

7.2 Non-verbatim exit interviews

Participant, Concentration:	**P75, ScB_7ten**
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• Stay perpendicular to FPSO to block ice • Stay North of lifeboat zone
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 2
3. What factors do you think were important for success in the scenario?	• Thrusters were used a lot. Very important to the scenario
4. What was the most challenging part of the scenario?	• Staying North was challenging due to pack ice build up and current
5. Would you change anything about your strategy/approach in the scenario?	• It might have been better to approach from the stern
6. Do you feel the decision support system adequately assisted you in the scenario?	• Yes
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 4
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• The physical reading was not specific enough (text) • Videos were the most informative
9. What would you have changed about the guidance from the decision support system in the scenario?	• More specific instructions
10. Other questions or comments.	• N/A

Participant, Concentration:	**T35, ScB_4ten**
Question	Reponses
1. Reflect on your performance in the scenario. What was your strategy?	• Asked for assistance from the start and followed it • Initial manoeuver had a positive effect on ice concentration • At the end, tried asking for assistance again. The new manoeuver recommended was less helpful
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 2
3. What factors do you think were important for success in the scenario?	• The (DSS) software. I would have been absolutely lost without it. • It was nice to get direction from the start
4. What was the most challenging part of the scenario?	• Understanding the ice dynamics is difficult
5. Would you change anything about your strategy/approach in the scenario?	• I would have left my position at the bow of the FPSO sooner and reproached the zone from the stern
6. Do you feel the decision support system adequately assisted you in the scenario?	• Yes.
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 3
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• The graphics (video) were the most helpful component. This was easier than reading the text based instructions on the fly. It would have been better for the video to do a continuous loop instead of stopping after 5 times.
9. What would you have changed about the guidance from the decision support system in the scenario?	• Being able to replay/loop the video. • Being able to reassess and get direction from current position and not from the start would have been better
10. Other questions or comments.	• Awesome.

Participant, Concentration:	**T70, ScB_7ten**
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• Tried to use the DSS for assistance. • Finished doing what it suggested and then did a loop around the vessel to do it again. • Pressed assistance once at the start
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 4

3. What factors do you think were important for success in the scenario?	• Getting distances from the FPSO was the most important • Practice during the scenario made me more comfortable near the end
4. What was the most challenging part of the scenario?	• Distances are hard to gauge • Familiarity with controls
5. Would you change anything about your strategy/approach in the scenario?	• I would have kept my distance between my vessel and the FPSO smaller if I did it again. By the end I was more confident to be close than at the beginning in my control abilities
6. Do you feel the decision support system adequately assisted you in the scenario?	• No. It was just a hint. I thought about my own solution in the minutes between briefing and beginning the scenario.
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 5. It was clear and straight forward
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• The video was the most useful.
9. What would you have changed about the guidance from the decision support system in the scenario?	• More colours for the instructions to help key words and instructions stand out. • Less words in the description in general
10. Other questions or comments.	• I did not read the instructions. Colours would catch the eye more in the instructions.

Participant, Concentration: W25, ScB_7ten	
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• Treat the simulated vessel as if it was a real vessel, accounting for current, wind, etc. • Was going to go around the FPSO in a circle originally but realized going in directly and using propwash while on a parallel course would be better
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 3
3. What factors do you think were important for success in the scenario?	• Tunnel thrusters were key • Stern station control was better for visibility with the propwash strategy • Using propwash was very effective
4. What was the most challenging part of the scenario?	• -Current. As soon as you clear the ice, more drifts in.
5. Would you change anything about your strategy/approach in the scenario?	• Not really, but a radar or AIS would be helpful

6. Do you feel the decision support system adequately assisted you in the scenario?	• Yes, because it gave me the most used way • Gave a scenario based opinion
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 4
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• The approach was the most helpful. Bird's eye view picture specifically, not the video • Only pressed assist once • Text was helpful
9. What would you have changed about the guidance from the decision support system in the scenario?	• Would not change anything
10. Other questions or comments.	• N/A

Participant, Concentration:	**C40, ScB_7ten**
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• Go down to the stern and then turn around and push ice up towards the front (J-approach) • DSS suggested a 3-point turn mid way through so that's what I tried after my first pass
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 3.4
3. What factors do you think were important for success in the scenario?	• My goal was to not hit the other boat and I succeeded in that
4. What was the most challenging part of the scenario?	• The ice just kept coming. It was so thick so you'd be pushing through it, then you'd hit a clear patch and all of a sudden would be going way too fast.
5. Would you change anything about your strategy/approach in the scenario?	• In my first pass, I would have done a three point turn and proceeded south the same way, instead of doing a full turn around the vessel like I did in my first pass.
6. Do you feel the decision support system adequately assisted you in the scenario?	• Yes
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 5
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• The advice to do a 3 point turn on the second pass and to get further ahead of the FPSO to prevent ice floating back down.

	• I liked how the DSS gave me a different scenario/advice on the second assistance request
9. What would you have changed about the guidance from the decision support system in the scenario?	• I wish it showed what I'd already done • Interface was good, but I never looked at the "suggested solution" numerical values • Did not use the text at all • Only looked at the videos and pictures
10. Other questions or comments.	• No. Pretty cool.

Participant, Concentration:	**V72, ScB_4ten**
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• Find direction of wind, current, how the ice was flowing when doing the scenario • Go in, hope for the best
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 3
3. What factors do you think were important for success in the scenario?	• Finding/being aware of the way the current was flowing
4. What was the most challenging part of the scenario?	• The presence of, trying to hold station
5. Would you change anything about your strategy/approach in the scenario?	• I would have come in half way through instead of the J-approach that the DSS suggested. This way, I'd go in directly and then hold the ice (Leeway)
6. Do you feel the decision support system adequately assisted you in the scenario?	• It did when I read the instructions. At first, no. I watched the video and started to try it, but it didn't really work. When I read the instructions near the end, I changed what I was doing to fit them and then it helped.
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 4.5
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• Text instruction 5 to use current was helpful. • Generally I'd rather watch something than read but the text instructions worked better than the video
9. What would you have changed about the guidance from the decision support system in the scenario?	• Make the video larger. Maybe have a monitor • Text was clear but I didn't read it until the end.
10. Other questions or comments.	• I enjoyed it

Participant, Concentration:	D84, ScB_7ten
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• I only have instructions, so try to watch the video, look at instructions and try to figure out what to do with my controls • I tried this assist a second time but it showed below average performance, but both suggested using leeway • I didn't know what pushing was, so I had to use intuition • My plan was to put the stern against the vessel and use leeway at the end
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 3
3. What factors do you think were important for success in the scenario?	• Understanding the instructions • Making a good initial approach and having a plane before entering the ice pack • It's tough to change once you're in the pack • I was worried about reversing in ice
4. What was the most challenging part of the scenario?	• See some of above • Current was difficult. Ice was constantly filling what you just cleared
5. Would you change anything about your strategy/approach in the scenario?	• I would try to come up from the stern and then use the leeway method
6. Do you feel the decision support system adequately assisted you in the scenario?	• Kind of. It gave me an idea of an approach but instructions weren't detailed enough • When you're in the thick of it, the lack of detail isn't there and it feels like the DSS drops the ball • I didn't know what the pushing technique was so that wasn't so helpful
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 3
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• The video was helpful • First instructions for approach are good, but the later instructions got me lost

9. What would you have changed about the guidance from the decision support system in the scenario?	• Having the video continue looping would be better • Having more details about techniques/ specific strategies
10. Other questions or comments.	• No

Participant, Concentration:	N17, ScB_4ten
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• J approach as I was told and tried to clear ice, but it kept coming. Tried to go back to hold it but it wasn't working. Requested help again and was told to leave, come back, and do a scenario where I am perpendicular to the FPSO
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 3
3. What factors do you think were important for success in the scenario?	• Being able to transfer between the bow and stern station • Having an initial course from the DSS
4. What was the most challenging part of the scenario?	• Being able to maintain a safe speed was difficult. There was so much to focus on.
5. Would you change anything about your strategy/approach in the scenario?	• Halfway through I was told to go out and do a 90 degree angle. I would have done that sooner if I did it again.
6. Do you feel the decision support system adequately assisted you in the scenario?	• Yes, especially in the beginning to see the video of how you should move to the vessel
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 4
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• Video was very beneficial • Heading and sped recommendations on the bottom left was helpful • I didn't use the text at all
9. What would you have changed about the guidance from the decision support system in the scenario?	• It would be beneficial if you didn't have to actively hit assist, if the assist was automatic. I pressed the button a lot.
10. Other questions or comments.	• No

Participant, Concentration:	B61, ScB_7ten
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• Go in, push ice forward of lifeboat station forward of FPSO, then come back and catch the stuff you missed. • I was too far forward though, which made it take longer than expected

2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 5
3. What factors do you think were important for success in the scenario?	• Understanding environmental conditions, understanding the use of prop-wash, understanding the use of the vessel as a plow • Understanding proper throttle inputs of the bow thrusters and stern thrusters (for a sideways push)
4. What was the most challenging part of the scenario?	• Visual reference • Being able to see all the growlers (ice floes). In real life, I could have someone on the bridge wing, or have someone calling distances for me from that position • I'm new in ice, so that was hard to know the behaviour of the floes.
5. Would you change anything about your strategy/approach in the scenario?	• I'd keep my strategy the same, but would change my approach further towards the stern. • Might even aim towards the lifeboat and then push up. This also gives you room and clear water to loop back down more easily
6. Do you feel the decision support system adequately assisted you in the scenario?	• Yes and no. No, because I wanted to do it myself. • Yes, in that it helped me verify/solidify my strategy since they were similar • It was like having an experienced master confirming that my approach was good. • IT gave some confidence and small adjustment tips, for example to push ice over the bow of the FPSO which I hadn't thought of
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 4
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• Push ice flow around the bow of the FPSO, and demonstrating the method to push the ice so that it bundles/clumps together. • The video was good. It would be better to combine that with step by step pictures, where you break the video into steps to show alongside
9. What would you have changed about the guidance from the decision support system in the scenario?	• When I'm in the situation and get advice as if I'm from the start, how can I adjust it to give me more current advice?

	• Currently it bases the advice on the original approach, not the current situation • Having the ability for the assist program to show next steps from multiple points would be helpful • I didn't look at the text • Instead of text, show each bullet point as a step by step picture. That much text is an overload
10. Other questions or comments.	• Text and numbers aren't super useful • The center of the interface draws the eye • Having a larger monitor, or a monitor from above would be better and less distracting. Looking down takes away the situational awareness

Participant, Concentration:	S63, ScB_4ten
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• I didn't push the ice under the lifeboat enough since I was scared to get too close to the FPSO. • Original plan was to push ice down, then return up and use leeway
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 3
3. What factors do you think were important for success in the scenario?	• Orientation of vessel and thrusters
4. What was the most challenging part of the scenario?	• Trying to get ice that was closer to the FPSO without knowing the instantaneous distance
5. Would you change anything about your strategy/approach in the scenario?	• I would try to get ice that was closer to the FPSO, use my thrusters more, and keep my speed down. • I tried to keep following the first one
6. Do you feel the decision support system adequately assisted you in the scenario?	• Yes, it was good. First strategy was the best, but the other two were below average. • I tried to follow the advice from the first one
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 4
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• The video was useful. I liked it a lot. I also liked the text steps and the end goals (suggested solution section), but mostly the video

9. What would you have changed about the guidance from the decision support system in the scenario?	• Having the video keep playing. • Step by step diagrams instead of just the picture to compliment instructions • It's currently hard to figure out where you are in the picture
10. Other questions or comments.	• Was really good

Participant, Concentration:	D86, ScB_7ten
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• Use/visua;ize myself on a compass rose from 0-360 degrees. The mast is the bow, and visualize reference points.
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 5
3. What factors do you think were important for success in the scenario?	• Knowledge of thrusters, speed you can maintain, how to use the wheel based on currents
4. What was the most challenging part of the scenario?	• Icebreaking. Ice floats as it breaks, and you're not stable, so keeping clear behind you was a challenge
5. Would you change anything about your strategy/approach in the scenario?	• I would go to the stern of the FPSO first, then come ahead with the wheel to starboard
6. Do you feel the decision support system adequately assisted you in the scenario?	• Yes, absolutely • You can see what other people did and see how it works. Not to copy them, but do something similar
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 5
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• Videos
9. What would you have changed about the guidance from the decision support system in the scenario?	• Video goes too quickly, and you have to focus on what you're doing • I didn't look at the text
10. Other questions or comments.	• No

Participant, Concentration:	L59, ScB_7ten
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• Go perpendicular to FPSO, block current of ice flowing down (with hull)

2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 4
3. What factors do you think were important for success in the scenario?	• DSS was very helpful. I had no idea what to do until I asked for assistance • Current was actually helpful • Thrusters were really helpful • Stern station was great for increased visibility
4. What was the most challenging part of the scenario?	• First, figuring out what to do and how to do it. Once I figured it out, stay in position and not letting ice past bow
5. Would you change anything about your strategy/approach in the scenario?	• I turned around to face FPSO stern first half way through due to the DSS suggestion, but maybe I would not have done that, or I would have approached stern first from the start. It look too much time turning around and let some ice in
6. Do you feel the decision support system adequately assisted you in the scenario?	• Yes, definitely
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 4.5
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• Positioning instructions (bottom left panel) with specific scenario instructions. The details from the "question marks" as well
9. What would you have changed about the guidance from the decision support system in the scenario?	• Video was too fast, and it stops after 6 times which isn't great • Didn't use picture
10. Other questions or comments.	• Don't think so, it was cool.

Participant, Concentration:	A92, ScB_7ten
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• I don't know… at first, trying to match the DSS. After, trying to go head on to the lifeboat, but it was moving slow and wasn't going where I wanted. • I wanted to be on a parallel course.
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 1.5
3. What factors do you think were important for success in the scenario?	• Knowing how to use the controls properly. I messed them up sometimes. • Thrusters confused me

	• Not being able to see the vessel (distances?)
4. What was the most challenging part of the scenario?	• It was going slower than expected and confused me • Tough to line up parallel to the vessel • Hard to do a sharp turn but had troubles with it
5. Would you change anything about your strategy/approach in the scenario?	• Would use the thrusters next time
6. Do you feel the decision support system adequately assisted you in the scenario?	• It was useful but I couldn't get the heading it suggested and I got impatient • I requested assistance 6 ish times
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 2.5
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• I was looking at the numbers (suggested solution) more than the pictures • The video was good • Text instructions were not as good as the video • Video was best
9. What would you have changed about the guidance from the decision support system in the scenario?	• Layout was 3 things (panels). Maybe have the video front and center with the text/instructions on the bottom • Less divided into three sections • Verbal instructions would be nice (text to speech) • More specifics for controls (more control focused) • You get lost in the strategy. More control specific would be good.
10. Other questions or comments.	• No. It was good. • I was scared to use the thrusters. Felt it was safer just to go straight and use the wheel. • Tough to visualize the ship and response is slow.

Participant, Concentration:	H29, ScB_4ten
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• I wanted to approach similar to the DSS, but use prop wash instead. • Manoeuver similar to the DSS because I thought it would put me in a better position

2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 4
3. What factors do you think were important for success in the scenario?	• Current and time management. The DSS showed me how much things change in 30 minutes because it shows the flow well.
4. What was the most challenging part of the scenario?	• Getting the initial clearing promptly. It takes some time for the area to get clear. • Maintaining the clear status was much easier.
5. Would you change anything about your strategy/approach in the scenario?	• Might have changed when I reversed to avoid backing into ice.
6. Do you feel the decision support system adequately assisted you in the scenario?	• Yeah, absolutely. Seeing the ice field moving was very helpful. • I didn't exactly do what it said, but seeing how it plays out was useful • Feels a bit panicked how fast the loop goes • Would have been helpful to see more failed runs.
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 5
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• I don't know of particular advice, but the visualization of drift
9. What would you have changed about the guidance from the decision support system in the scenario?	• The scenario with the looping vessel (W28) was weird. It never suggested a different scenario. Remove bad scenarios.
10. Other questions or comments.	• No

Participant, Concentration:	E90, ScB_7ten
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• Initially, clear by pushing • Saw DSS said perpendicular, so switched to that approach • Use Prop Wash to clear a path. Reverse and then position to use leeway
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 3
3. What factors do you think were important for success in the scenario?	• Current, and thrusters were helpful • Habituation was helpful
4. What was the most challenging part of the scenario?	• Making a plan

5. Would you change anything about your strategy/approach in the scenario?	• Go with forward clearing instead of clearing from reverse (with propwash) • I forgot to watch my speed
6. Do you feel the decision support system adequately assisted you in the scenario?	• Yes, I guess medium • I forgot to look at it after the initial • I pressed it and forgot/didin't fully understand
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 3
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• End picture. I never would have thought of holding station and letting the current do the work
9. What would you have changed about the guidance from the decision support system in the scenario?	• Would be better for people who know nautical language • "Safe distance" is vague
10. Other questions or comments.	• no

Participant, Concentration:	F41, ScB_4ten
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• Use assist button, suggested going to stern and using thrusters to push ice
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 4
3. What factors do you think were important for success in the scenario?	• Current was a challenge
4. What was the most challenging part of the scenario?	• Current
5. Would you change anything about your strategy/approach in the scenario?	• Yes, I trued to turn to be parallel to FPSO at the end. Should have done this earlier
6. Do you feel the decision support system adequately assisted you in the scenario?	• Yes. Quite good
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 5
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• Using thrusters .I pressed it 5 times. • Video was the best
9. What would you have changed about the guidance from the decision support system in the scenario?	• At the beginning, it showed what I wanted. • I didn't like that it showed something new when I continued pressing the button
10. Other questions or comments.	• No

Participant, Concentration:	R12_ScB_4ten

Question	Reponses
1. Reflect on your performance in the scenario. What was your strategy?	• Go to stern, then move ice in forward direction • Move to bow of FPSO and play with ice in that region
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 5
3. What factors do you think were important for success in the scenario?	• Directions. Bow thrusters and positioning of vessel. Once in position is correct, moving ice is easy • Thrusters
4. What was the most challenging part of the scenario?	• Trying to clear ice in bow section
5. Would you change anything about your strategy/approach in the scenario?	• Change the approach by moving it from perpendicular to FPSO, to pushing with side from stern
6. Do you feel the decision support system adequately assisted you in the scenario?	• Yes, it was helpful
7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 3
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• Video helped me a lot
9. What would you have changed about the guidance from the decision support system in the scenario?	• I was trying to see previous suggestions, but you can't go backwards • It only shows new suggestions
10. Other questions or comments.	• No

Participant, Concentration:	O52, ScB_7ten
Question	**Reponses**
1. Reflect on your performance in the scenario. What was your strategy?	• Try to follow DSS, but ability to manoeuver was impacted by lack of shiphandling knowledge
2. Rate your overall performance in completing the scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 2
3. What factors do you think were important for success in the scenario?	• Experience with vessel
4. What was the most challenging part of the scenario?	• Manoeuvring at low speed (heading specifically) • Managing rudder and thrusters
5. Would you change anything about your strategy/approach in the scenario?	• Overall, no. More familiarity and practice would help me execute
6. Do you feel the decision support system adequately assisted you in the scenario?	• Yes, but there are improvements that could be made

7. Rate the performance of the decision support system for helping you complete this scenario. (1 is not very successful, 3 is somewhat successful, 5 is very successful)	• 3
8. What advice or instructions did you find the most useful from decision support system in the scenario?	• Advice at the outset on general strategy
9. What would you have changed about the guidance from the decision support system in the scenario?	• Make it so that ships path line is easier to read. Currently, it's only one pixel against a background. • Make it so that you can slow the video down • Make it so that the assist starts at the current position
10. Other questions or comments.	•

7.3 Ethics

7.3.1 Ethics Approval Letter

MEMORIAL UNIVERSITY

Interdisciplinary Committee on Ethics in Human Research (ICEHR)

St. John's, NL Canada A1C 5S7
Tel: 709 864-2561 icehr@mun.ca
www.mun.ca/research/ethics/humans/icehr

ICEHR Number:	**20220529-EN**
Approval Period:	July 28, 2021 – July 31, 2022
Funding Source:	NSERC [RGCS# 20151916]
Responsible Faculty:	Dr. Brian Veitch Faculty of Engineering and Applied Science
Title of Project:	*Evaluating efficacy of a decision support system (DSS) for emergency ice management in a simulated environment*

July 28, 2021

Mr. Jonathan Soper
Faculty of Engineering and Applied Science
Memorial University of Newfoundland

Dear Mr. Soper:

Thank you for your submission to the Interdisciplinary Committee on Ethics in Human Research (ICEHR), seeking ethical clearance for your research project. The Committee appreciates the care and diligence with which you prepared your application. The project is consistent with the guidelines of the *Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans* (TCPS2). *Full ethics clearance* is granted for **one year** from the date of this letter. ICEHR approval applies to the ethical acceptability of the research, as per Article 6.3 of the *TCPS2* (2018). Researchers are responsible for adherence to any other relevant University policies and/or funded or non-funded agreements that may be associated with the project. If funding is obtained subsequent to ethics approval, you must submit a Funding and/or Partner Change Request to ICEHR so that this ethics clearance can be linked to your award.

The *TCPS2* **requires** that you **strictly adhere to the protocol and documents as last reviewed** by ICEHR. If you need to make additions and/or modifications, you must submit an Amendment Request with a description of these changes, for the Committee's review of potential ethical issues, before they may be implemented. Submit a Personnel Change Form to add or remove project team members and/or research staff. Also, to inform ICEHR of any unanticipated occurrences, an Adverse Event Report must be submitted with an indication of how the unexpected event may affect the continuation of the project.

The *TCPS2* **requires** that you submit an Annual Update to ICEHR before **July 31, 2022**. If you plan to continue the project, you need to request renewal of your ethics clearance and include a brief summary on the progress of your research. When the project no longer involves contact with human participants, is completed and/or terminated, you are required to provide an annual update with a brief final summary and your file will be closed. All post-approval ICEHR event forms noted above must be submitted by selecting the ***Applications: Post-Review*** link on your Researcher Portal homepage. We wish you success with your research.

Yours sincerely,

Russell J. Adams, Ph.D.
Chair, Interdisciplinary Committee on
Ethics in Human Research

RA/bc

copy: Supervisor – Dr. Brian Veitch, Faculty of Engineering and Applied Science
Director, Research Grant and Contract Services

7.3.2 Ethics Revision approval Letter

MEMORIAL UNIVERSITY

Interdisciplinary Committee on Ethics in Human Research (ICEHR)

St. John's, NL Canada A1C 5S7
Tel: 709 864-2561 icehr@mun.ca
www.mun.ca/research/ethics/humans/icehr

ICEHR Number:	**20220482-EN**
Approval Period:	July 20, 2021 – July 31, 2022
Funding Source:	RGCS#: 20151916 & 20210835; Veitch
Responsible Faculty:	Dr. Brian Veitch Faculty of Engineering and Applied Science
Title of Project:	*Evaluating efficacy of a decision support system (DSS) for emergency ice management in a simulated environment*
Amendment #:	***02***
Title of Parent Project:	*Evaluation of the effects of digital decision support technology on marine ice management performance in a simulator environment*
ICEHR Number:	*20211198-EN*

February 8, 2022

Mr. Jonathan Soper
Faculty of Engineering and Applied Science
Memorial University

Dear Mr. Soper:

The Interdisciplinary Committee on Ethics in Human Research (ICEHR) has reviewed the proposed revisions for the above referenced project, as outlined in your amendment request dated January 14, 2022. We are pleased to give approval to the recruitment methods and revised experience questionnaire, as described in your request and subsequent communication, provided all other previously approved protocols are followed. In addition, researchers are responsible for adherence to the **MUN COVID-19 guidelines** (https://www.mun.ca/covid19/researchers/index.php).

The *TCPS2* **requires** that you **strictly adhere to the protocol and documents as last reviewed** by ICEHR. If you need to make any other additions and/or modifications during the conduct of the research, you must submit an Amendment Request with a description of these changes, for the Committee's review of potential ethical issues, before they may be implemented. Submit a Personnel Change Form to add or remove project team members and/or research staff. Also, to inform ICEHR of any unanticipated occurrences, an Adverse Event Report must be submitted with an indication of how the unexpected event may affect the continuation of the project.

Your ethics clearance for this project expires **July 31, 2022**, before which time you must submit an Annual Update to ICEHR, as required by the *TCPS2*. If you plan to continue the project, you need to request renewal of your ethics clearance, and include a brief summary on the progress of your research. When the project no longer requires contact with human participants, is completed and/or terminated, you need to provide an annual update with a brief final summary, and your file will be closed.

All post-approval ICEHR event forms noted above must be submitted by selecting the *Applications: Post-Review* link on your Researcher Portal homepage.

The Committee would like to thank you for the update on your proposal and we wish you well with your research.

Yours sincerely,

James R Drover

James Drover, Ph.D.
Vice-Chair, ICEHR

JD/bc
cc: Supervisor – Dr. Brian Veitch, Faculty of Engineering and Applied Science

7.3.3 Informed Consent Form

Informed Consent Form

Title:	Evaluation of the effects of digital decision support technology on marine ice management performance in a simulator environment
Researcher:	Jonathan Soper, Graduate Student, Faculty of Engineering and Applied Science, Memorial University, (613) 306-7284, jksoper@mun.ca
Supervisor:	Dr. Brian Veitch, Supervisor, Faculty of engineering and Applied science, Memorial University, (709) 864-8970, bveitch@mun.ca

You are invited to take part in a research project entitled *"Evaluation of the effects of digital decision support technology on marine ice management performance in a simulator environment."*

This form is part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. It also describes your right to withdraw from the study. In order to decide whether you wish to participate in this research study, you should understand enough about its risks and benefits to be able to make an informed decision. This is the informed consent process. Take time to read this carefully and to understand the information given to you. Please contact the researcher coordinator *Jonathan Soper*, if you have any questions about the study or would like more information before you consent.

It is entirely up to you to decide whether to take part in this research. If you choose not to take part in this research or if you decide to withdraw from the research once it has started, there will be no negative consequences for you, now or in the future.

Introduction:

I am Jonathan Soper, a Masters Student in the Faculty of Engineering and Applied Science's Department of Ocean and Naval Architectural Engineering at Memorial University of Newfoundland in St. John's. As part of my Masters thesis I am conducting research under the supervision of Dr. Brian Veitch. The research is being conducted as a part of the Safety at Sea project, funded by Husky Energy, Natural Sciences and Engineering Research Council (NSERC), and Virtual Marine

Purpose of Study:

In previous studies from the Safety at Sea group, experienced seafarers and cadets were evaluated in the ice management simulator to set a baseline of performance for a given experience level. It was found that on average, experienced seafarers performed better than cadets. A second study examined the use of simulator training on inexperienced cadets and found that performance could be significantly improved through the use of training to a level where cadets were able to perform at a similar level to the experienced seafarers. A third study was completed to gather expert knowledge from experienced seafarers which could then be used in the construction of a digital decision support system (DSS). You are being asked to be a participant on a study designed to evaluate the efficacy of this decision support technology to see whether it can improve cadet performance to a level which equals performance with training. This research may be used to inform future requirements for decision support systems in seagoing ships and can guide the implementation of further improvements to the next iteration of DSS.

What You Will Do in this Study:

If you choose to participate in this study, you will be asked to complete a simulated emergency ice management procedure in an ice management simulator.

You will work with a member of the research team to schedule times that are convenient for you to participate in this study. It is expected that your session for this study will take a maximum of three (3) hours.

Each session will take place at the Safety at Sea project's Simulation Lab (EN1035) in the Engineering and Applied Sciences (SJ Carew) building on Memorial University's St. John's campus.

You will arrive at the ice management simulator at the scheduled time where you will meet a member(s) of the research team.

The sessions will be split into four parts: (1) Briefing, (2) Familiarization Trial, (3) Ice Management Scenario, and (4) Feedback and Closing.

Refreshments (water and snacks) will be on hand for you during the trials. We will have time for you to take multiple breaks throughout the sessions to allow you to have some refreshments, move around outside of the simulator, or use the washroom.

1. *Briefing:*

We will explain the research and an opportunity to ask questions or express concerns. If satisfied, you will indicate your free and informed consent by completing this Informed Consent Form.

Before you start any trials, we will ask you to complete an experience questionnaire. We will also ask you to fill out a simulator sickness questionnaire (SSQ) in order for us to establish a baseline score for you. We will administer the SSQ to you throughout the trials to see if you are developing simulator sickness, which will be indicated by a higher score.

2. *Familiarization Trial:*

Once in position on the console, you will be asked to perform the 3-familiarization trials. These trials are designed to allow you to get familiar with the ice management simulator, and how the ship handles in the simulation. These trials are expected to take approximately 5-10 minutes for a total of 15-30 minutes.

A final familiarization trial will be undertaken to ensure that you understand how to use the DSS. You will be shown how to ask for assistance, and how to interpret the advice, data, and visual images given.

You will also be informed that the DSS provides support, but that it is not mandatory to follow the advice should you feel it is unhelpful.

After the familiarization trials are completed, we will move on to the ice management scenario.

3. *Ice Management Scenario:*

The ice management scenario consists of a 30 minute simulated emergency response using the DSS, followed by a debriefing.

Prior to starting the ice management scenario, you fill out a SSQ and go through a planning exercise with us.

When the scenario has been completed, you will be escorted off the ice management simulator and fill out another SSQ to determine if you are experiencing any symptoms of simulator sickness. You will then be shown a sped up video replay of your current scenario, where we will ask you interview style questions about your ice management techniques. We will ask you a series of questions to get your opinion on your performance and what factors you considered during ice management. You will be asked whether you used the DSS, what advice it gave, and whether you found it to be helpful. Your answers will be recorded on paper.

4. *Feedback and Closing:*

You will be asked to give feedback on the habituation scenarios, the emergency scenario, and post-trial questions. After this, the session will be completed.

Length of Time:

Your session is expected to take a maximum of three (3) hours.

Your Participation:

Your involvement in this study is voluntary and confidential. As such, you participation is not a requirement of your employment. We will not identify you as a participant in this study, nor report your participation to your superiors or co-workers. The data collected in this study will not be traced back to you.

Withdrawal from the Study:

You can withdraw from this study at any point during your participation without giving any reason, and all data collected up until that point will be destroyed. There are no consequences to you for withdrawal from the study. If you choose to withdraw from the study after your participation, your data can be removed from the study up to two weeks after your participation. To withdraw from the study just inform the Principal Investigator, Jonathan Soper

Compensation and Travel Expenses Reimbursement:

Possible Benefits:

There will be no direct benefit to you for participating in this study.

Data collected from this study will benefit in the development of decision support systems for use by navigation crew on ships.

Possible Risks:

A risk associated with participating in this study is the potential development of simulator-induced sickness. Simulator-induced sickness is very similar to motion sickness and can occur when people use equipment such as virtual reality headsets or simulators. Symptoms can include fatigue, headache, eye strain, difficulty focusing, increased salivation, sweating, nausea, stomach awareness, blurred vision, dizziness, vertigo, and burping. The symptoms can sometimes occur during, immediately after or several hours after exposure to the simulator.

We will be monitoring you for simulator sickness throughout the ice management scenarios by asking you complete the simulator sickness questionnaire (SSQ). If you self-report any of the above symptoms as “moderate” or “severe”, we will pause the trials and you will be provided with a rest period until your symptoms have subsided. You can decide whether you would like to resume the trials after the rest period. If the symptoms subside, and you choose to do so, we can continue with the trials. If you choose to not continue with the trials, we will stop the trials and you will exit the simulator.

If after the session ends the symptoms of simulator sickness persist for more than 20 minutes, we will arrange for you to get home safely.

Your performance in the simulator will be recorded throughout the study. For some individuals, this may cause performance anxiety or stress. This anxiety or stress may be caused by poor performance in the scenarios, by the difficulty or novelty of the task, or by repeated trials. To reduce the likelihood of anxiety and stress, where possible, we will guide you through the scenarios of the study. You will receive a break between scenarios to rest and you will be instructed not to worry or dwell on the previous scenarios.

You will be reminded that if you are not comfortable with any aspect of the trials, then you have the right to withdraw from the study at any point. To reduce the likelihood of embarrassment, you will perform the task individually and you will be reminded that your performance in the simulator will be anonymous.

That is, your data is not linked to your identity and that your performance or withdrawal will not be reported to anyone.

If at any time you experience symptoms or discomfort, which prevent you from continuing in this study you retain the right to withdraw from the study.

As discussed in the *Anonymity* section of this form, the researchers cannot guarantee your complete anonymity in this research. While your name will not be reported, you may be identifiable to other people based on other information you provide. This means there is a risk of being identified based on your participation in this study. To reduce the likelihood of you being identified the researchers will avoid reporting any identifiable information such as specific vessels you have worked on.

There is a risk of embarrassment in this study if you feel you cannot answer the researchers' questions adequately. To reduce the likelihood of embarrassment you will be reminded that you are not being tested by these simulator trials.

Confidentiality:

The ethical duty of confidentiality includes safeguarding participants' identities, personal information, and data from unauthorized access, use, or disclosure. Protecting your privacy and maintaining confidentiality is important to the research team. The information gathered in this study will be used solely for research purposes. Only researchers involved in this study will have access to the data.

Anonymity:

Anonymity refers to protecting participants' identifying characteristics, such as name or description of physical appearance.

Protecting your privacy is an important goal for the research team and this means ensuring all personal data recorded during your participation remains anonymous. You will not be directly identified in publications. The study will use a number to identify you, not your name. For example, researchers will use an alphanumerical participant code (e.g. AB001) to identify you in all reports of your data including when direct quotations are used. Only the principal investigator will be able to link this number to your name. Measures have been taken to remove any other possible identifiers other than your name, like number of years of experience onboard a specific type of vessel, for instance. You will not be video or audio recorded in this study

Recording of Data:

As part of this study, we will be collecting the following data from you:

- Name and contact information.
- Shipboard experience.
- Simulator sickness questionnaire scores.
- Ice management scenario performance (from the simulator)
- Post-trial debrief questionnaire.

Use, Access, Ownership, and Storage of Data:

The research team will collect and use only the information they need for this research study. Your name and contact information will be kept in a locked office on a password protected computer by the research team at MUN. It will not be shared with others without your permission. You will receive a randomized alphanumeric participant code (e.g. AB001). All information collected from you will be recorded with the participant code. Your name will not appear in any report or article published as a result of this study.

Information collected, anonymized, and used by the research team will be stored by the Principal Investigator, Jonathan Soper.

A hardcopy of your questionnaire responses will be kept in a filing cabinet in a locked office accessible by the research team. This data will have no identifiable information and will be kept separate from your signed consent form. Electronic data recorded in this study will be kept in a password protected file on a hard drive accessible only by the research team. This data will not have any identifiable information. Data will be kept for a minimum of five years, as required by Memorial University's policy on Integrity in Scholarly Research. After five years, all electronic records of your participation will be permanently deleted and all paper files will be appropriately destroyed. Data collected in this study will be documented in an Ocean Engineering Research Center (OERC) report. This will make the data accessible to other researchers but not the general public. This report will not include any of your identifiable information.

Reporting of Results:

The research team intends to publish the findings of this study in peer-reviewed journals and academic conferences. Formal reports will be made available to the research project partners (the National Research Council, Husky Energy, and Virtual Marine). Upon completion, my Masters thesis will be available at Memorial University's Queen Elizabeth II library, and can be accessed online at: http://collections.mun.ca/cdm/search/collection/theses. The data will be reported in a summarized statistical and descriptive form. Individual information or data will not be reported without your exclusive written consent.

Sharing of Results with Participants:

When data analysis is completed a report will be prepared and participants who wish to be informed of the results will have the opportunity to receive a copy of this report. The results will also be reported in my Masters thesis, which be available at Memorial University's Queen Elizabeth II library, and can be accessed online at: http://collections.mun.ca/cdm/search/collection/theses

Questions:

You are welcome to ask questions before, during, or after your participation in this research. If you would like more information about this study, please contact: Jonathan Soper (jksoper@mun.ca) or Brian Veitch (bveitch@mun.ca).

ICEHR Approval Statement:

The proposal for this research has been reviewed by the Interdisciplinary Committee on Ethics in Human Research and found to be in compliance with Memorial University's ethics policy. If you have ethical concerns about the research, such as the way you have been treated or your rights as a participant, you may contact the Chairperson of the ICEHR at icehr@mun.ca or by telephone at 709-864-2861.

NRC Research Ethics Board Statement:

This study has also been approved by the NRC Research Ethics Board (NRC-REB) under protocol number 2019-119. REB review seeks to ensure that research projects involving humans as participants meet Canadian standards of ethics. Any questions or concerns about the ethics of this study may be directed to the NRC-REB Secretariat, NRC-REB@nrc-cnrc.gc.ca, (613) 949-8681.

Consent:

Your signature on this form means that:

- You have read the information about the research.
- You have been able to ask questions about this study.
- You are satisfied with the answers to all your questions.
- You understand what the study is about and what you will be doing.
- You understand that you are free to withdraw participation in the study without having to give a reason, and that doing so will not affect you now or in the future.
- You understand that if you choose to end your participation during data collection, any data collected from you up to that point will be destroyed.
- You understand that if you choose to withdraw after data collection has ended, your data can be removed from the study up to two weeks after your participation.

By signing this form, you do not give up your legal rights and do not release the researchers from their professional responsibilities.

Your Signature Confirms:

☐ I have read what this study is about and understood the risks and benefits. I have had adequate time to think about this and had the opportunity to ask questions and my questions have been answered.

☐ I agree to participate in the research project understanding the risks and contributions of my participation, that my participation is voluntary, and that I may end my participation.

☐ A copy of this Informed Consent Form has been given to me for my records.

________________________________ ________________________________

Signature of Participant Date

Researcher's Signature:

I have explained this study to the best of my ability. I invited questions and gave answers. I believe that the participant fully understands what is involved in being in the study, any potential risks of the study and that they have freely chosen to be in the study.

________________________________ ________________________________

Signature of Principal Investigator Date

Participant Number: _______________
Date: _______________

7.3.4 Simulator Sickness Questionnaire

Participant Number: ______________ Date: ______________ Time: ______________

When: ☐ Before Habituation 1 ☐ After Habituation 3

☐ After Scenario 1 ☐ After Scenario 2 ☐ After Scenario 3

☐ After Scenario 4 ☐ After Scenario 5

----------------------------PLEASE DO NOT WRITE ABOVE THIS LINE-----------------------------

Simulator Sickness Questionnaire

Please indicate the severity of symptoms that apply to you right now. Also note that there is no obligation to answer any or all questions if you do not wish to do so, but you must answer all questions in order to continue the study. There are no consequences for withdrawal from the study.

Symptom	**None (0)**	**Mild (1)**	**Moderate (2)**	**Severe (3)**
General discomfort				
Fatigue				
Headache				
Eyestrain				
Difficulty focusing				
Increased salivation				
Sweating				
Nausea				
Difficulty concentrating				
Fullness of head				
Blurred vision				
Dizziness (with eyes open)				
Dizziness (with eyes closed)				
Vertigo				
Stomach awareness				
Burping				

Kennedy, R. S., Lane, N. E., Berebaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology*, 3(3), 203-220.

7.3.5 Experience questionnaire

Experience Questionnaire

Please answer the following questions but feel free to omit any that you do not wish to answer. If something is unclear, please ask the research coordinator. Your answers are confidential.

Question	Answer
1. What is your year of birth?	______________________
2. What is your gender?	☐ Male ☐ Female ☐ Non-binary ☐ Prefer not to say Self-identify: ________________
3. In what year of study are you enrolled?	☐ 1st year ☐ 2nd year ☐ 3rd year ☐ 4th year ☐ Over 4th year
4. Are you enrolled in a nautical science program?	☐ Yes ☐ No
5. What academic program are you enrolled in?	______________________
6. Approximately how many months experience do you have at sea?	______________________
7. On what types of vessels have you operated? (Select all that apply)	☐ OSV / AHTS ☐ Icebreaker ☐ Tanker / Bulk / Cargo ☐ Ferry / Coastal ☐ I have not spent time at sea
8. Have you ever operated in sea ice?	☐ Yes ☐ No

9. What types of operations did you perform while in ice? (Select all that apply)	☐ Watchkeeping during transit ☐ Manoeuvring ship while being escorted ☐ Manoeuvring ship to escort another vessel ☐ Ice management (open water) ☐ Ice management (confined water) ☐ Towing or emergency response ☐ I have only observed operations in ice ☐ I have not operated in ice

Participant Number: _______________

10. Where have you obtained your experience in operating in ice? (Select all that apply)	☐ Great lakes ☐ Gulf of St. Lawrence ☐ Coastal Newfoundland and Labrador ☐ Arctic (north of 60) ☐ Baltic Sea ☐ Caspian Sea ☐ Sea of Okhotsk ☐ Antarctic ☐ I have not operated in ice
11. Approximately how many years have you spent in the presence of sea ice?	________________________
12. What types of shore-based training have you taken for operating in ice? (Select all that apply)	☐ Basic training in ice operations ☐ Advanced training in ice operations ☐ Attendance at professional seminars discussing techniques and procedures relevant to ice operations ☐ I have never received training related to ice operations
13. Do you have any experience using a marine simulator? (Select all that apply)	☐ Training for navigation in open water ☐ Training for navigation in ice ☐ Research study ☐ I have no experience using a marine simulator

7.3.6 Recruitment Document

The text below was circulated to students within the Marine Institute's Nautical Science program. The same text was also printed and posted to the walls of the Marine Institute in the form of a poster.

Recruitment Email: Volunteers Needed to Operate a Marine Simulator

Researchers at Memorial University are studying the effect of decision support guidance on ice management performance in a simulator environment. The outcomes of the research will help to develop early stage decision support technology and could inform future ways of providing onboard training.

The Experiment:

- This research is being completed as part of a Master's Degree in Ocean and Naval Architectural Engineering under the supervision of Dr. Brian Veitch at Memorial University of Newfoundland.
- The research is being conducted using the Ice Management Simulator located in the Engineering and Applied Sciences (SJ Carew) building on Memorial University's St. John's campus.
- If you participate, you will be asked to attend 1 session (which could take up to 3-4 hours to complete).
- Refreshments and breaks will be provided.
- Exposure to simulators has been known to cause simulator-induced sickness. Researchers will monitor participants throughout the study for symptoms of simulator-induced sickness.
- Volunteers can withdraw from the study at any time and for any reason. There are no consequences for withdrawal from the study.

Who can participate?

- Post secondary students
- Ages 18 years of age or older
- Must have normal or corrected-to-normal vision
- Must have no prior experience with the MUN Ice Management Simulator

Recruitment is open from September 2021 – April 2022.

All participants will receive a $20 gift card for participating.

Reimbursement can be provided for travel to/from Memorial University (by taxi or other) up to $100.

Participation in this study is not a program requirement, will not effect student grades, and will not be reported to instructors, other students, or school administrators.

To protect participants and the research team during the experimental session, All participants will be required to follow the COVID-19 Health and Safety plan. This plan will be emailed to all interested volunteers. Please note that before coming to campus for the experimental session you must review the COVID-19 health and safety guidance document, and complete the COVID-19 Daily Self-Assessment Tool (available online at http://www.mun.ca/covid19/faculty-staff/COVID19_self_assessment.pdf)

If you are interested or have any questions, please contact:
Jonathan Soper
Email: jksoper@mun.ca
Phone: (613) 306-7284

If you know anyone why may be interested in this study, please give them a copy of this information

The proposal for this research has been reviewed by the Interdisciplinary Committee on Ethics in Human Research and found to be in compliance with Memorial University's ethics policy. If you have ethical concerns about the research, such as your rights as a participant, you may contact the Chairperson of the ICEHR at icehr.chair@mun.ca or by telephone at 709-864-2861