Simulator Training for Lifeboat Maneuvers

Lochlan E Magee
Kingston, Ontario, Canada
lochlanmagee@yahoo.com

Jennifer J E Smith
St. John's, Newfoundland, Canada
jenniferjesmith@gmail.com

Randy Billard, Anthony Patterson
Virtual Marine Technology Inc.
St. John's, Newfoundland, Canada
firstname.lastname@vmtechnology.ca

ABSTRACT

Lifeboats are essential life-saving equipment for many commercial (and military) vessels and offshore platforms. Their coxswains usually learn to maneuver on calm waters at a nautical school for safety reasons. Simulator training could avoid the environmental constraints that limit training with an actual lifeboat (e.g., frozen water) while enlarging the experience (e.g., high winds). However, the agencies that regulate training for lifeboats have little available evidence of simulator effectiveness since little research has been conducted in this domain (Power-Macdonald, MacKinnon, Simões Ré, Power & Baker, 2011).

As a first step in an ongoing investigation of simulator training for emergency operations of a lifeboat, we conducted a behavioral experiment to assess transfer of training of maneuvers practiced in a simulator to control of a lifeboat in benign environmental conditions. To reduce the effects that natural abilities and prior experience could have on the experimental outcome, all participants initially attempted a basic maneuvering course with the lifeboat. They were then paired by their performance. One member of each pair, chosen at random, subsequently received simulator training. After all members of the simulator trained group were trained to criterion on a virtual representation of the test course, the members of both groups individually completed the test course to criterion on the water with the actual lifeboat. The test course involved a sequence of challenging maneuvers. Ten participants in each group completed the test course.

In this paper, we report learning in the simulator, and we compare initial transfer of training and improvement with practice to criterion by each treatment group. A transfer effectiveness ratio (TER) is calculated. In addition, we report the effect of practicing to criterion within the simulator on training transfer. We discuss the implications of the results for lifeboat training regulators, as well as their relevance to the broader training community.

ABOUT THE AUTHORS

Lochlan E Magee is an independent consultant. He previously conducted and managed the research, development and evaluation of training simulators for Defence Research and Development Canada (DRDC) for more than three decades. He performed investigations of training effectiveness in maritime, land and air domains and human factors studies. He represented Canada at international meetings of The Technical Cooperation Panel (TTCP) and NATO. Immediately prior to retirement, he conducted an evaluation of a simulation for training emergency procedures aboard a Victoria Class submarine. He was an Adjunct Professor in the Math and Computer Sciences Department of the Royal Military College of Canada from 2008 to 2014. Dr. Magee possesses a PhD in Experimental Psychology from the University of Toronto. His area of study was human information processing, particularly human perception.

Jennifer J E Smith is an independent research consultant and Human Factors Research Coordinator for the Safety at Sea Program at Memorial University. Her role involves the design, implementation and analysis of human factors experiments with marine simulation prototypes. Jennifer has worked on various projects at Memorial for the past seven years. Prior to her role as Human Factors Coordinator, she was the Project Manager of two marine simulation based research projects. She is currently a candidate for a Ph.D. in Engineering at Memorial. She possesses a Master's degree in Engineering from Memorial and holds a Bachelor's degree in Applied Science from Acadia University and a Bachelor's degree in Biological Engineering from Dalhousie University. Her research interests are in Human Factors, Marine Simulation, and Offshore Safety.

Randy Billard is Chief Technical Officer and Executive Vice-President of Virtual Marine Technology Inc. (VMT). He is responsible for leading VMT's team of engineers and computer scientists in the continuous development of small craft simulation technologies. He has expertise in military and civilian simulator development and project management and is actively involved in ocean engineering and simulation R&D communities. He has experience in managing technical product development from early stage prototypes to commercially ready systems which includes validation and testing of simulators at various technology readiness levels. He is a P. Eng, holds a Bachelor's degree in Mechanical Engineering and a Master's Degree in Ocean and Naval Architectural Engineering from Memorial University of Newfoundland, with expertise in vessel and wave modeling. He is currently a candidate for a Ph.D. in Engineering at Memorial University. To avoid conflict of interest, Mr Billard did not participate in the collection or analysis of the experimental data.

Captain Anthony Patterson is a Master Mariner and possesses a Bachelor of Maritime Studies from Memorial University of Newfoundland. His main area of expertise is marine operations, especially under emergency and/or harsh environments. Currently the President and CEO of Virtual Marine Technology Inc. (VMT), he has conducted and managed research programs related to Search & Rescue and has developed, delivered and evaluated training programs for the Canadian Coast Guard, the Marine Institute of Memorial University of Newfoundland, and major oil companies. His research experience includes the development of computerized search planning techniques and the evaluation of human factors associated with marine operations in harsh environments. His training experience includes the development of marine emergency management courses, oil spill response courses, lifeboat evacuation courses and the Canadian Pleasure Craft Operator Certificate program. He is currently the National Education Chair for the Company of Master Mariners of Canada and is a member of the Canadian delegation to the International Maritime Organization (a legislative body of the United Nations). To avoid conflict of interest, Captain Patterson did not participate in the collection or analysis of the experimental data.

Simulator Training for Lifeboat Maneuvers

Lochlan E Magee
Kingston, Ontario, Canada
lochlanmagee@yahoo.com

Jennifer J E Smith
St. John's, Newfoundland, Canada
jenniferjesmith@gmail.com

Randy Billard, Anthony Patterson
Virtual Marine Technology Inc.
St. John's, Newfoundland, Canada
firstname.lastname@vmtechnology.ca

INTRODUCTION

Lifeboats are essential life-saving equipment for many types of vessels and offshore platforms, such as oil rigs. Although their purpose is for an emergency escape, possibly in rough seas, their coxswains usually learn to maneuver the lifeboat on calm waters at a nautical school to avoid risks during training. Simulator-based training could avoid risks during training and provide relief from the environmental constraints (e.g., frozen water) that limit training with an actual lifeboat. Simulator-based training could also afford a more varied training experience by exposing the student to the effects of high winds, large waves, or poor visibility within the safety of a virtual environment. However, the agencies that regulate training for lifeboats have little performance-based evidence of simulator training effectiveness, since little research has been conducted in this domain (Power-Macdonald, MacKinnon, Simões Ré, Power & Baker, 2011). Consequently, regulators lack behavioral evidence that could help inform a decision to authorize the use of simulators for training lifeboat maneuvering skills. A decision about the use of simulators for training lifeboat maneuvers is important to the offshore oil industry and others who want to improve the quality, cost-effectiveness and availability of training.[1]

This paper reports a behavioral experiment that was conducted to assess transfer of training from a lifeboat simulator to the control of a real lifeboat. It is part of an on-going investigation of emergency skills training for lifeboat operations in the harsh environments representative of offshore worksites. The basic method of assessing training transfer is to compare the performance of a simulator trained group to the performance of a control group that did not receive prior simulator training when both groups are required to make use of the operational equipment (Blaiews, Puig & Regan, 1973). Power-Macdonald et al. (2011) employed this method to validate the use of a simulator for training control of a lifeboat among obstacles on the water that represented a patch of floating ice. They found evidence of positive transfer. In our experiment, we evaluate a range of competencies fundamental to the proficient control of a lifeboat using a test course that presented a sequence of challenging tasks as they might be encountered in natural circumstances. Since there are risks in using a lifeboat on the open water, we conducted the experiment under benign environmental conditions to see if the simulator could substitute for an actual lifeboat, which is restricted in use for training to these conditions.

Although the specific operational context for our investigation is lifeboat maneuvers, we provide behavioral results of relevance to the broader simulator training community. We provide evidence of training transfer effectiveness and results indicating that the instructional practice of training to criterion in the simulator can be detrimental to training transfer. Since users can feel that they are in the virtual environment of the simulator, while they are physically elsewhere in the real world, and because this phenomenon, called presence (Witmer & Singer, 1998), is thought to facilitate experiential learning and training transfer, we also investigate its relationship to training transfer.

[1] Subsequent to the experimental work reported here, some regulators have accepted simulator training for lifeboat maneuvers.

METHODS

Task

The on-water test course was designed by two subject matter experts (SMEs). It is illustrated in Figure 1. The course was designed to model the tasks and challenges that a lifeboat coxswain might encounter when abandoning a platform or ship such as navigating through a debris field, steering to a rendezvous point, recovering a person in the water and coming alongside another vessel to transfer personnel. The course requires coxswains to conduct several subtasks performed in sequence. The subtasks include steering to a landmark, maneuvering through a slalom course (consisting of 3 gates), maneuvering alongside a vessel (simulated by two marks tied together), steering by compass, recovering a Person in the Water (PIW) and coming to a stop between two marks. This set of subtasks permits the assessment of a number of competency measures including steering by sight and compass, maintaining a steady course, approaching, stopping at, and maneuvering around various anchored and drifting objects.

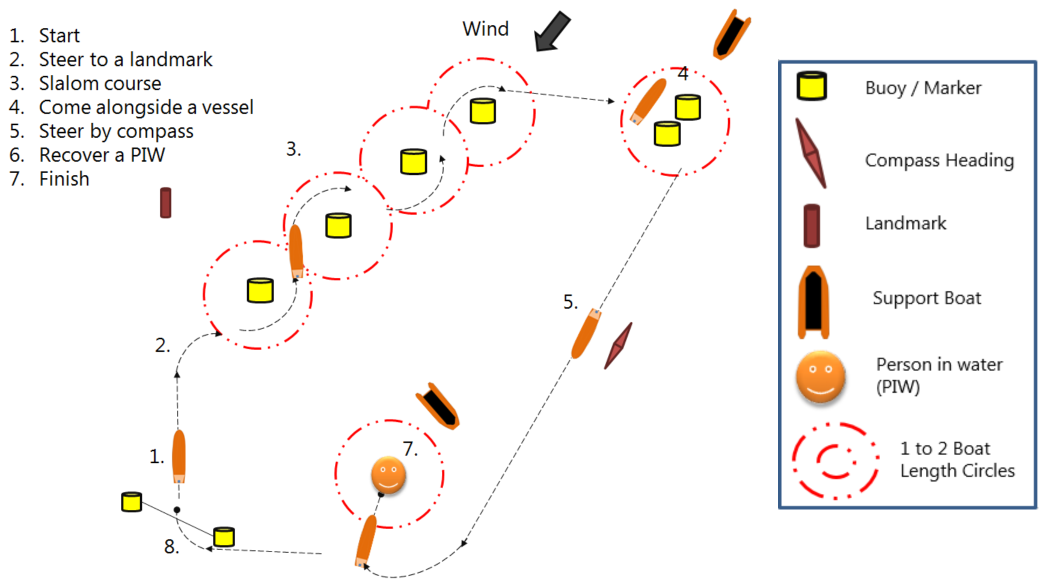

Figure 1. On-water Evaluation Course

The course was set on the waters of North Arm at Holyrood, Newfoundland, Canada. This site was used previously by MacDonald et al. (2011) and was chosen for on-water testing because it is sheltered from the winds and waves of the North Atlantic. The actual path length of the course was approximately 2.0 km (1.1 NM), allowing a completion time of about 18 minutes (ranging from 12 to 25 minutes). The main axis of the slalom task was pointed into the wind. The course was set each day, and sometimes reset several times during the day, if the prevailing wind direction changed by more than 10 degrees. Testing on the water was only conducted if the water was hazard free, the wind was 15 knots or less, if the wave height was less than 0.5 m, if the ambient air temperature was greater than -10 degrees C, and if visibility was greater than one nautical mile.

An onboard instructor/operator (I/O) managed and assessed each participant. In order to complete the course successfully, the participant needed to perform all of its subtasks successfully. The participant needed to achieve performance criteria set for each subtask while demonstrating good coordination of the throttle and steering wheel and account of obstacles, wind and waves. For example, in coming alongside another vessel, the participant needed to select the safest approach plan given the environmental circumstances, combine wheel and throttle control appropriately during the approach, stop the lifeboat within 38m of the marks representing the other vessel, and hold position without drift of the lifeboat by more than a boat length for at least 10 seconds. The scoring criteria for all the subtasks were developed by the I/O to reflect a standard of proficiency that would be accepted by the lifeboat training community. To aid consistent application of the scoring criteria, a paper-based rubric and score sheet were used and all assessments were made by the same I/O. The I/O was provided no information about the participant's treatment group.

Participants

Twenty-four (24) healthy volunteers between the ages of 18 and 65 were recruited. Only 20 of the participants completed the study. Four were unable to complete the study due to excessive delays caused by weather and scheduling conflicts; their data were excluded from our analyses. The participants' ages ranged from 21 to 52 years with a median age of 33.5 years. Six of the participants were female. All were physically healthy, and if they were using medication, their condition needed to be stable. All participants were required to have normal or corrected to normal (20/20) vision and all needed to be able to swim. No recruit was allowed to participate if they had prior lifeboat training experience or anticipated receiving lifeboat training elsewhere during the course of this experiment.

All participants were requested to refrain from consuming alcohol 24 hours before any training or test event. All participants were briefed on the risks, benefits and responsibilities associated with the experiment. All participants were required to provide written consent to participate and were informed that they had the right to withdraw from the experiment at any time without explanation or prejudice. In addition, they were told how they would be assigned to treatment groups. The participants were compensated for travel expenses to and from the test facilities and were provided an incentive of $50 for each session that they attended.

If the participant did not possess a Pleasure Craft Operator Card (PCOC), the participant needed to complete the test necessary to obtain a PCOC prior to their operation of the lifeboat. This accreditation is a legal requirement of Transport Canada, and it is necessary for the operation of small boats in Canadian waters. Four of the participants possessed a PCOC beforehand. One other possessed a bridge watch keeping certificate in addition to the PCOC. All others completed an on-line course.

The experimental protocol for this study was approved by the National Research Council of Canada Research Ethics Board and by the Health Research Ethics Board of Newfoundland and Labrador.

Lifeboat

A side view of the lifeboat that we used for this study is shown in Figure 2. This lifeboat is approximately 9.4 m long, 3.5 m wide and 6 m high, with a draft of 2.9 m. Its empty weight is approximately 5806 kg. Fully loaded, it carries up to 72 people and weighs about 11506 kg. In our experiment the lifeboat was empty except for the participant and the I/O. The coxswain controls the lifeboat from within the raised cockpit at the stern of the lifeboat. There are two forward windows, one on the port and another on the starboard side of the lifeboat.

Figure 2. Lifeboat Underway with Hatches Open

Simulator

Det Norske Veritas Germanischer Lloyd (DNV-GL, 2014) provides standards for maritime simulator systems and issues certificates of compliance declaring the technology suitable for use in specified marine training programs. The certificate confirms the functional fidelity of the simulator and its mathematical modeling. The simulator that was used in this study has all the features of a class S (special tasks) simulator, which has been certified by DNV-GL to create realistic situations for the competencies that we trained and tested. A picture of the simulator is shown in Figure 3. The simulator was made by Virtual Marine Technology Inc. (VMT) of St. John's, Newfoundland. It provided a physical mockup of the cockpit of the lifeboat and allowed the participants to operate the controls needed to maneuver a virtual lifeboat within a virtual environment. Liquid Crystal Display (LCD) displays behind the windows of the cockpit presented computer-generated scenes of the water surface, the sky, and physical objects in the surrounding visual environment, including visible parts of the simulated lifeboat, surrounding land forms, and floating objects as they would be seen through the windows of an actual lifeboat at the seated position of the coxswain. The hydrodynamic model of the simulated lifeboat was matched to the handling characteristics of the actual lifeboat. Sound effects, such as modulation of the engine noise, were provided to correspond with the sounds of the real environment. No physical motions were imparted to the participants within the simulator. The simulator was located in a room at VMT that was free of distractions. An employee of VMT acted as the instructor-operator (I/O). The I/O controlled and monitored the simulator during each training session.

Figure 3. Lifeboat Simulator

Experimental Design

A between groups, repeated measures experimental design was used (2 groups, 6 trials maximum). The experiment was conducted in three phases as follow:

Phase 1: Familiarization and Assessment

This phase of the experiment had two purposes. One was to familiarize all participants with the physical environment of the lifeboat, instructions for its operation, and some hands on experience on the water. The other purpose was to assess the relevant skills and abilities of the participants so that we could create two groups of initial equal capability to avoid the effects that prior experience or inherent abilities could have upon the experiment outcome. All participants received an abbreviated coxswain training course that involved about three hours of classroom lecture including portions of an Advanced Coxswain Training (ACT) course developed by VMT. The participants received this training in groups of three or more and were shown a diagram of the assessment course illustrated in Figure 4, which was designed by two SMEs to introduce and assess basic lifeboat handling skills. The participants were informed that this course was composed of several tasks that would be used to assess their ability to navigate to a landmark, approach a mark, stop at a mark, navigate using a compass, maneuver around a mark, approach a vessel (represented by two static marks), come alongside a stopped vessel, approach the finish marks, stop between the finish marks, and move forward at no wake and at consistent speed.

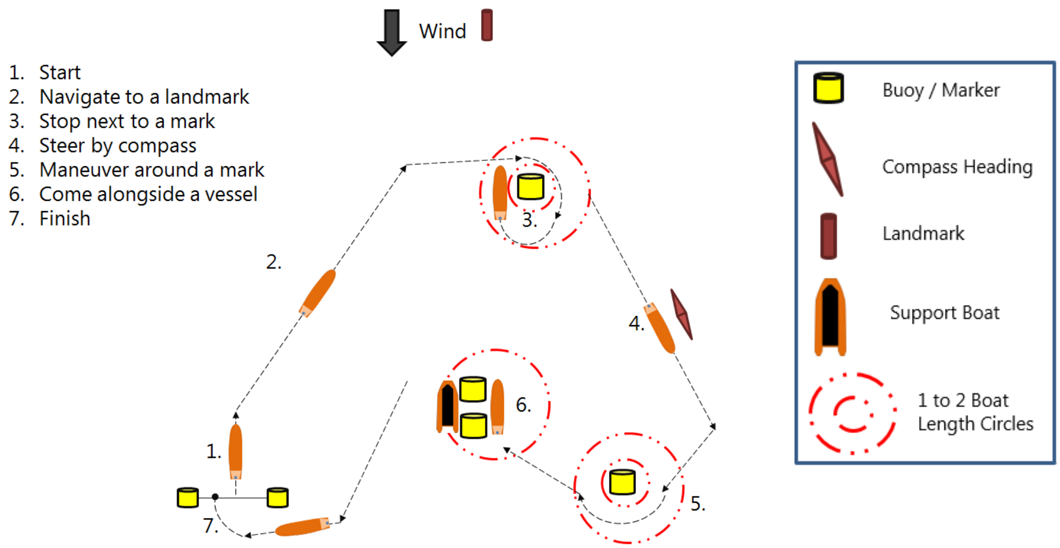

Figure 4. On-water Assessment Course

Testing on the water was conducted individually several days after the classroom training. The delay between classroom training and on-water assessments varied due to weather and scheduling constraints, averaging 6 days, with a range of 2 to 11 days. Each participant was given the opportunity to complete a simple loop with the lifeboat before the start of the assessment course. Each participant was made aware that their performance would be assessed and they were informed about the assessment criteria (as described previously). The participants were also told that the I/O would not answer any questions about the control of the lifeboat, but that he would provide situational awareness (e.g., distance and bearing to marks). The participants were required to wear an approved Personal Floatation Device (PDF), an immersion suit, a hardhat and earplugs. The experimenter was prepared to terminate the session if there was a violation of safety conditions. As a safety precaution, the speed of the lifeboat was not allowed to exceed 8 knots when it was operated by the participants. The I/O used a rubric and scoring scheme to assess the participants. The same I/O assessed all participants.

Phase 2: Training

The scores obtained by the participants on the assessment course were ranked and then used to assign them to one of the two treatment groups. Individuals of close rank were paired. One member of the pair was placed in one group and the other member of the pair was placed in the other group to counterbalance abilities. One of the groups was then picked randomly as the experimental (E) group for simulator training. The members of this group received practice in the simulator as soon as possible after Phase 1. The average delay was 10 days; the range was between 4 and 15 days. The control (C) group did not receive additional practice of any kind.

The members of the E group maneuvered the simulated lifeboat around a virtual test course identical to the evaluation course illustrated in Figure 1. The simulated wind speed was set to 9 knots, which was the average wind speed recorded during the assessment trials of Phase 1. The direction of the simulated wind was in line with the main axis of the slalom course. The visibility within the virtual environment was slightly overcast. The participants were encouraged to achieve the criteria set for all subtasks on each attempt of the virtual course so that they could complete it successfully as a sequence from beginning to end. The average amount of time required to complete one attempt at the virtual course was 23 minutes, ranging from 14 to 37 minutes. The cumulative amount of time taken by each participant to achieve criterion on the virtual course was 111 minutes, ranging from 61 to 177 minutes.

The I/O used the rubric and score sheet to provide feedback at the end of each trial, but did not coach the participant or provide guidance during or after any trial. All feedback was outcome based so that the training would be experiential and uninfluenced by the instructional skill of the I/O. Two different VMT employees were needed to be the I/O during the simulator training sessions in order to minimize the delays between phases of the experiment. The I/O allowed brief rests between trials to avoid unwanted side effects. As many as 9 trials were allowed to reach criteria. Once this limit was reached, the training session was concluded.

An experimenter observed all interactions between the I/O and the participant to assure conformity with the protocol. The experimenter also observed the participant for signs and symptoms of simulator-induced sickness (SIS). If the experimenter observed signs or symptoms of SIS, such as yawning, pallor, sweating, burping, or loss of balance, or if the participant reported illness, a rest period was scheduled and the participant was asked to complete a SIS questionnaire (Kennedy, Lane, Berebaum, & Lilenthal, 1993).

Four of the 10 participants indicated 2 to 5 symptoms of SIS, all rated as minimal. The symptoms included some of the following: nausea, stomach awareness, burping, general discomfort, fatigue, eyestrain, sweating, difficulty focusing, difficulty concentrating, fullness of head and dizziness with eyes open. One of the four participants who reported minimal symptoms also rated some symptoms as moderate. Symptoms of fatigue, fullness of head and dizziness (eyes closed) were rated as moderate by this person. The onset of the symptoms rated minimal occurred after 5 unsuccessful attempts to complete the course, and the moderate symptoms occurred after 8 unsuccessful attempts. Due to these reports, the experimenter withdrew the participant from further training and reminded the participant about the potential hazards of SIS and ways to reduce the after effects. The experimenter contacted the participant by phone later in the day to find that the participant had recovered.

Presence

A questionnaire developed by Witmer & Singer (1998) was used to assess the extent to which the participants in the E group felt that they were present with the virtual environment of the simulator. The questionnaire was given to each member of the E group immediately after they completed the simulator training session.

Phase 3: Evaluation

The members of the C group were tested on the evaluation course between 7 and 34 days after initial training, with an average delay of 19 days. The evaluation trials for the experimental group occurred between 15 and 29 days, with an average delay of 22 days, after initial training. This delay included a delay of 12 days on average between simulator training and evaluation, with a range of 3 to 19 days. Most of the delays were due to foul weather at the test site and the remaining delays were due to scheduling conflicts. In order to limit the effects that environment conditions could have on performance, the experimenter attempted to alternate testing of the members of the C and E groups as much as possible. To limit the potential effects that the I/O could have on performance, no information about the participant's prior training was given to the I/O, and the participants were discouraged from revealing their prior training to the I/O. All evaluation trials were conducted individually, and no participant had an opportunity to observe the performance of others. Each participant was reminded of the task, the objectives, safety precautions and right to withdraw. Each participant attempted the evaluation course until they reached criterion, or until 6 unsuccessful attempts at the course were made.

RESULTS & DISCUSSION

Training with the Simulator

Figure 5 provides a plot of the cumulative number of trials taken by members of the E group to achieve criterion on the task with the simulator. This figure shows that several attempts at the task were required to achieve criterion and that the performance of the group improved with practice until trial number 7, after which a plateau seems to be reached. This plateau is due to one participant who failed to reach criterion within the nine attempts that were allowed. The median number of trials required to achieve criterion on the simulated evaluation course is 4.5.

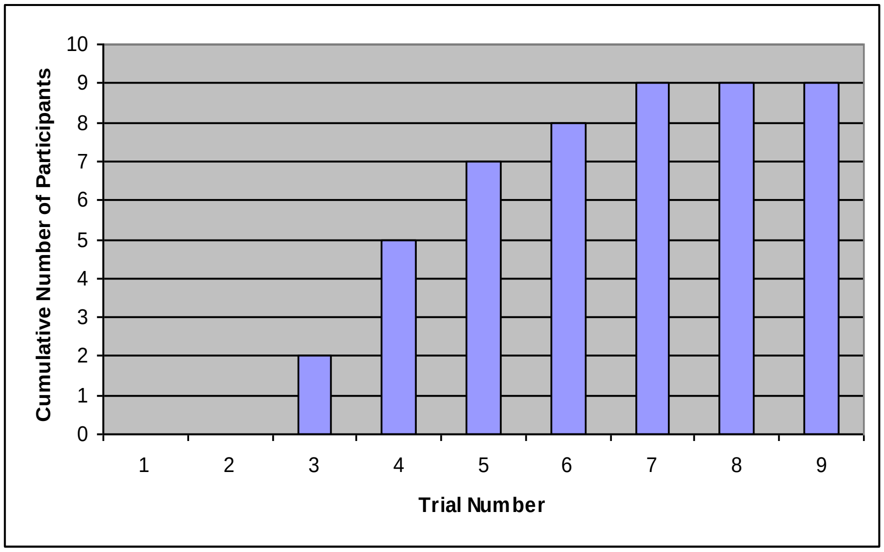

Figure 5. Cumulative Number of Participants to Achieve Criterion by Trial with the Simulator

Initial Transfer of Training

Evidence of positive training transfer from the simulator to the lifeboat is shown for Trial 1 in Figure 6, which provides a bar graph that cumulates and compares the number of participants in each treatment group that achieved criterion with repeated attempts at the evaluation course.

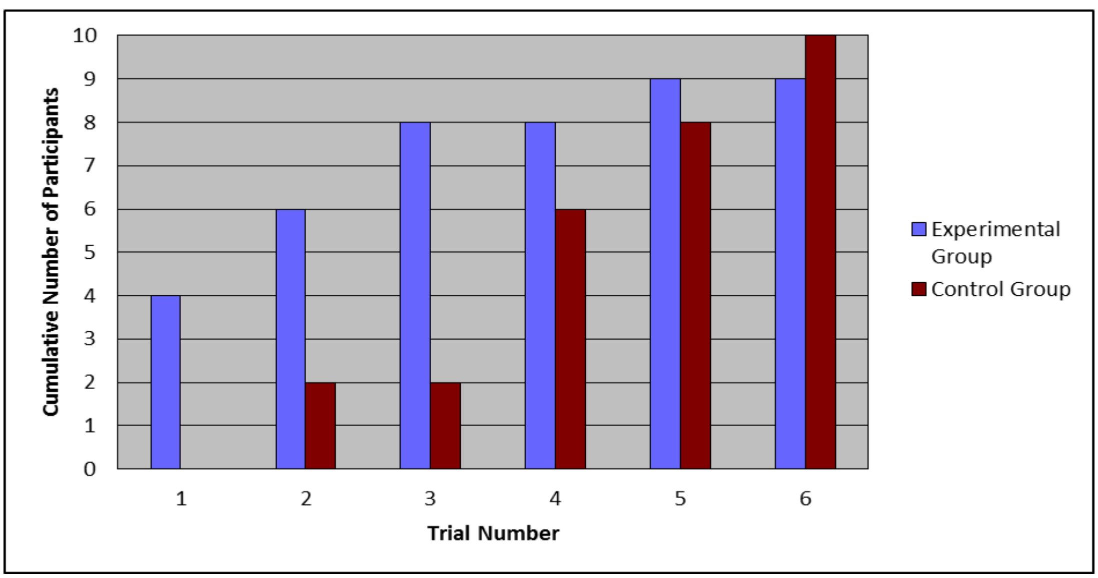

Figure 6. Cumulative Number of Participants to Achieve Criterion by Trial with the Lifeboat

The plot for the E group shows that 4 of its 10 members achieved criterion on their first attempt at the course. In comparison, the plot for the C group shows that no member of this group achieved criterion on their first attempt. The Fisher Exact Test statistic for this comparison of counts is $p < 0.10$, which indicates that the result is not likely due to chance and that it can be attributed to positive transfer of training from the simulator to the lifeboat. It is important to note that even though 9 of the 10 participants in the E group previously achieved criterion in the simulator, six did not achieve criterion immediately on the water. Perhaps surprising is the finding that the one participant, who failed to achieve criterion in the simulator after 9 attempts, achieved criterion immediately when tested on the water. These findings prompted an examination of the effect that the amount of practice in the simulator had on initial transfer. We found that all of participants who achieved initial success on the water trial had made five or more attempts in the simulator and that only one participant failed to achieve success on the water if they attempted the course five or more times in the simulator. In comparison, no participant achieved initial success on the water if they previously made less than five attempts in the simulator. The Fisher Exact Test of these data indicates that an outcome of this type or one more extreme has a probability of less than 5% occurrence by chance.

The influence that simulator training has on initial performance on the water is revealed more dramatically when the scores of each group are compared. The average score obtained by the C group on their first attempt at the evaluation course was 34.0, out of a possible 60. In comparison, the average scored obtained by the simulator trained group was 49.2. A statistical test (one-tailed t-test) of this difference for independent means reveals that the outcome is unlikely due to chance; the value of $t(18) = 3.8$, $p \leq 0.007$. Therefore, we can conclude that there is reliable, positive training transfer from the simulator to the real world and that the simulator is a valid training device for teaching lifeboat maneuvering skills.

Transfer of Training with Practice

The number of trials taken to achieve criterion on the test course is our principal measure of training transfer. This measure accounts for adaptation to the change in the task environment, and it provides the essential information to calculate a transfer effectiveness ratio (TER). The TER also considers the amount of practice in the simulator when evaluating transfer (Blaiwes et al, 1973). Positive transfer of training from the simulator to the lifeboat is seen in a comparison of the number of trails taken by each group to achieve criterion. The median number of trials taken by the simulator trained participants was 2, whereas the median number of trials for the control group was 4. A Mann-Whitney U test (one way) of the statistical significance of the difference in the median number of trials taken by the members of the two groups to reach criterion gives a U-value of 23, which is less than the critical value of 27 for significance of $p \leq 0.05$. This means that practice with the simulator reliably reduced the number of trials required to achieve training criterion on the water. This result confirms positive training transfer from the simulator to the real world and the validity of the simulator for training lifeboat maneuvering skills.

A TER is a useful measure of training transfer since it represents the operational savings as a proportion of simulator exposure (Povenmire and Roscoe, 1971). In other words, it tells us how much experience with the real system can be saved with simulator experience (Fletcher, 2013). It can be measured by time, trials, or errors to a criterion (Blaiwes et al., 1973).

The TER based on the trials to criterion can be computed as follows.

$$TER = \frac{Tc - Te}{Ts} \qquad (1)$$

where,

Tc is the median number of trials taken to reach criterion by the C Group
Te is the median number of trials taken to reach criterion by the E Group
Ts is the median number of trials taken to reach criterion by the E Group in the simulator

If we substitute the values that we obtained in this experiment for the variables in this equation, then TER = 0.44. This means that one trial on the water can be saved with 2 or 3 practice attempts with the simulator.

Presence

It is important to note that the answers provided by the participants to the presence questionnaire (PQ) were informed by their prior experience with the actual lifeboat on the water. From them, we learned that the participants felt the following: 1. that they gained proficiency in maneuvering the simulated lifeboat, 2. that they were engaged and in control of the simulation, and 3. that they were able to concentrate on the task without being distracted by limitations of the sensory cueing or control systems. No relationship was found between the participants' PQ scores and the number of trials that they required to achieve criterion with the simulator, $r_{xy} = 0.14$, $p > 0.10$, or with the lifeboat on water, $r_{xy} = 0.16$, $p > 0.10$. On this basis we can conclude that PQ score is not a useful predictor of performance with the simulator, nor is it a useful predictor of simulator training transfer to the lifeboat.

CONCLUSIONS

The results of this experiment provide positive, performance-based evidence of the effectiveness of the simulator for training coxswains to maneuver a lifeboat in benign conditions. Our conclusion is substantiated by a reliable difference between the median number of trials required to achieve performance on a test course by the simulator trained group and the median number of trials required by a control group that did not receive simulator training. Our conclusion is applicable to a broad range of potential trainees since the age, gender, English language skills and prior abilities of the participants in this experiment were wide ranging. Our finding is consistent with the finding of Macdonald et al. (2011), who assessed a lifeboat simulator for training maneuvers through an ice pack, and Magee (1997), who assessed a virtual reality simulator for training officers-of-the-watch to maneuver a minesweeper (a relatively small ship) in formation with other minesweepers.

Our experiment was designed to capture behavioral data in an ecologically valid context. We evaluated the performance on open water, with an actual lifeboat, using a test course that involved obstacle avoidance, precise maneuvering, approach and rounding of marks, approach and stopping alongside marks representing a stationary vessel, steerage by visual landmark, steerage by compass and pick up of a person in the water (simulated). Each attempt at the test lasted about 18 minutes and was conducted as a sequence of tasks that could be encountered in natural circumstances. These features of the test speak to the ecological credibility of the results. However, the training and testing were conducted under benign environmental conditions for safety reasons. Consequently, generalization of the results to more hazardous operational conditions should be made with caution, but not dismissed, since simulators are often used for training hazardous tasks in stressful flight environments, such as aircraft carrier landings at night, helicopter deck landings in rough seas, or space shuttle maneuvers. Hence, we have no reason to believe that lifeboat simulator training could not be extended to hazardous conditions, while affording a cost-effective and safe learning environment.

We calculated the transfer effectiveness ratio (TER) of the simulator to be 0.44, which indicates that 2 or 3 practice attempts in the simulator are equivalent to one practice attempt on the water. Since four or more practice attempts on the water were needed by 50% of the members in the control group to achieve criterion on the water, this finding indicates an opportunity for time and cost savings with simulator training in addition to the conveniences of simulator training and risk avoidance. It is important to note that our results are based on experiential learning with the simulator and the lifeboat. Feedback was provided to the participants about their performance on each trial, but the I/O did not guide or advise the participants on how to improve their performance. Instruction is an important factor in learning (Salas, Rosen, Held & Weissmuller, 2009) and an instructor operator station (IOS) can offer many advantages for delivering simulator-based training (Walwanis, Swanson, & Atkinson, 2013). Since the IOS capabilities of the simulator used in this experiment were not exploited by an instructor, we can predict that the amount of transfer could be larger. The results also indicate that the amount of transfer from the simulator to the lifeboat could be greater if more practice is given. By training to criterion, early success in the simulator meant fewer practice attempts in the simulator, which, in turn, seems to have led to poorer performance on the water. The participants who practiced most in the simulator tended to do best on the water, whether they reached criterion in the simulator or not. The most likely explanation for this positive effect is overlearning. Since we required the participants in the simulator trained group to achieve criterion on all performance measures of the virtual test course, some task elements of the course would have been overlearned in the simulator, especially by the participants who required many attempts to complete the course from beginning to end because only one or a few task elements prevented them from achieving early success. In a review of overlearning for training, Driskell, Willis & Copper (1992) report that adding 50% more trials, beyond the number of trials needed to reach criterion, produces a reliable, positive benefit for subsequent performance and that additional overlearning produces greater benefits. In other words, retention is generally much better when practice continues after a task is learned. This learning principle seems to be evident in the results that we obtained for the simulator trained group.

Although subjective ratings of presence within the virtual environment of the simulator did not predict training effectiveness, the participants felt confident about their training in the simulator. User confidence is an important indicator of user acceptance and for this reason alone, positive ratings are a welcome find for a training simulator. McCauley (2006) has noted that the concept of presence is similar to the concept of perceptual fidelity, which is a subjective impression of the realism of a simulator, and that the relationship between perceptual fidelity and performance in a simulator has not been established. In other words, the relationship between the fidelity of a simulator and its training effectiveness is not certain (Fletcher, 2013). The effects of presence are also known to be inconsistent (Sadowski & Stanney, 2002). We found no relationship between PQ scores and performance in the simulator. Since subjective ratings of presence in the lifeboat simulator do not predict success with the lifeboat, the combined results imply that the amount of practice obtained in the simulator is the overriding predictor of success. This conclusion questions the value of instructional strategies that train to criterion, or rely on apparent fidelity or presence, as a predictor of success. It encourages a strategy of overlearning with a simulator, which is an instructional approach applicable to a wide range of simulators and training applications.

ACKNOWLEDGEMENTS

We very much appreciate the help of the staff of VMT who performed the roles of instructor/operator. In particular, we thank Mr. Ryan Kelly, Director of Curriculum Development, who helped design the assessment and evaluation courses and, single handedly, developed and delivered all classroom training, developed the performance metrics, maintained and operated the lifeboat, coordinated all safety personnel and conducted all on-water trials and assessments of the participants' performance. We also thank Petroleum Research Newfoundland & Labrador and the Industrial Research Assistance Program of the National Research Council of Canada as sponsors.

REFERENCES

Blaiews, A. S., Puig, J. A., & Regan, J. J. (1973). Transfer of training and the measurement of training effectiveness. *Human Factors*, *15*, 523-533.

Det Norske Veritas Germanischer Lloyd. (2014). *Maritime simulator systems. Standard DNVGL-ST-0033:2014-08.* Retrieved April 26, 2016, from https://rules.dnvgl.com/docs/pdf/DNVGL/ST/2014-08/DNVGL-ST-0033.pdf.

Driskell, J.E., Willis, R.P., & Copper, C. (1992). Effect of overlearning on retention. *Journal of Applied Psychology*, 77 (5), 615-622.

Fletcher, J.D. (2013). Economic considerations in simulation-based training. In C. Best, G. Galanis, J. Kerry & R. Sottilare (Eds.), *Fundamental Issues in Defence Training and Simulation*. (pp. 135-154). Surrey, England: Ashgate Publishing Limited.

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.

McCauley, M. E. (2006). *Do army helicopter training simulators need motion bases?* (Technical Report 1176). Arlington, VA: United States Army Research Institute for the Behavioral and Social Sciences.

Magee, L. E. (1997). Virtual reality simulator (VRS) for training ship handling skills. In R.J. Seidel & P.R. Chatelier (Eds.), *Virtual Reality: Training's Future?* (pp. 19-29). London: Plenum Press.

Povenmire, H. K., & Roscoe, S. N. (1971). An evaluation of ground-based flight trainers in routine primary flight training. *Human Factors, 13*, 109-116.

Power-MacDonald, S.P., MacKinnon, S., Simões Ré, A., Power, J., & Baker, A. (2011). *Effects of simulator training on novice operators' performance in simulated ice covered waters.* (Technical Report TR-2011-15). St. John's, NL: National Research Council Canada – Institute for Ocean Technology.

Sadowski, W., & Staney, K. (2002). Presence in virtual environments. In K. M. Stanney (Ed.), *Handbook of virtual environments*. (pp 791-806). New Jersey: Erlbaum.

Salas, E., Rosen, M.A., Held, J.D., & Weissmuller, J.J. (2009). Performance measurement in simulation-based training: A review and best practices. *Simulation & Gaming*, *40*(3), 328-376.

Walwanis, M. M., Swanson, B.L., & Atkinson, B.F.W. (2013). The role of instructor operator stations in training. In C. Best, G. Galanis, J. Kerry & R. Sottilare (Eds.), *Fundamental Issues in Defence Training and Simulation.* (pp. 179-191). Surrey, England: Ashgate Publishing Limited.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments, 7*(3), 225-240.