Using Simulation to Assess Performance in Emergency Lifeboat Launches

Randy Billard
Virtual Marine
St. John's, NL, Canada
randy.billard@virtualmarine.ca

Jennifer Smith
Independent Research Consultant
St. John's, NL, Canada
jenniferjesmith@gmail.com

ABSTRACT

Launching a lifeboat in an emergency requires safety-critical proficiency which can only be achieved and maintained with hands-on practice. Simulators have been specifically created for offshore oil and gas personnel to practice lifeboat launching and maneuvering using representative equipment and virtual environments. As an alternative to live boat training, lifeboat simulators allow for practice in plausible, high-risk events in a safe, realistic environment. An automated simulator is an alternative offering the benefit of on-demand practice while expanding training capabilities. Providing training for these types of scenarios presents challenges for evaluating trainee performance in conditions traditionally not used in training because performance metrics may not exist. The study uses simulation to assess performance in lifeboat training from two perspectives: a live instructor and automated simulator. An experiment was performed to evaluate performance of lifeboat operators in an emergency scenario which included adverse weather and hazards. A simulator was used to provide a safe and controlled means to assess trainee performance. A rubric was created to define scoring for launching and maneuvering tasks in weather, including moderate sea sates. The rubric identified quantitative measures which could be used by the simulator and live instructor to assess performance. The study compared performance measurements taken by a live instructor and simulator with automated tracking as each assessed participants in a simulated emergency exercise. The results show the simulator provided an advantage of being able to consistently track performance on tasks where multiple performance criteria were measured simultaneously. The study also identified limitations in the simulator which were not present in instructor led evaluations, including subjective measures made through visual observation. The paper discusses how simulation can be used to automate scoring and reduce instructor workload, and how simulators can be used to measure trainee preparedness for an emergency event with waves and hazards.

ABOUT THE AUTHORS

Randy Billard is Chief Technical Officer and Executive Vice-President of Virtual Marine. He is responsible for leading Virtual Marine's team of engineers in the continuous development of simulation technologies. Randy has expertise in military and civilian simulator development and project management and is actively involved in ocean engineering and simulation R&D communities. 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, with a research concentration in measuring human performance in simulation-based training programs.

Jennifer J E Smith is a part-time research consultant working with Virtual Marine on a study to evaluate training methods for emergency lifeboat operation procedures. Jennifer is also a Human Factors Research Coordinator for the Safety at Sea Program at Memorial University. Over the past 7 years she has been involved in the design, implementation and analysis of human factors experiments with marine simulation prototypes designed for offshore safety training. 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 a Bachelor's degree in Applied Science from Acadia University and a Bachelor's degree in Biological Engineering from Dalhousie University.

Using Simulation to Assess Performance in Emergency Lifeboat Launches

Randy Billard
Virtual Marine
St. John's, NL, Canada
randy.billard@virtualmarine.ca

Jennifer Smith
Independent Research Consultant
St. John's, NL, Canada
jenniferjesmith@gmail.com

BACKGROUND

Lifeboat operators are required to have essential skills needed to launch a lifeboat in an emergency. As the launch of a lifeboat is not a routine event, coxswains are required to practice regularly to maintain the requisite skills. Lifeboat coxswains are expected to be able to launch and maneuver a lifeboat in environmental conditions that prevail in their location of operation, which commonly include wind speeds above 20 knots and wave heights from 3-5 meters if operating in the North Atlantic Ocean (C-Core, 2015). Coxswains are also expected to know operating procedures of launching a lifeboat, performing emergency procedures while in the water, and dealing with environmental hazards. Lifeboat coxswains typically complete initial training at an onsite facility and then perform regular recurrency training to maintain their skills. Recurrency training is normally conducted using drills involving a real lifeboat which include the launch of the lifeboat and performing simple maneuvering tasks on the water. This training is normally performed in benign weather conditions to minimize risk to trainees.

A new training alternative has been the use of simulation allowing for practice in controlled environments using virtual cues and representative lifeboat equipment. Simulation allows for training to be extended to conditions which are plausible in an emergency, including weather with moderate to high sea states and reduced visibility. Hazards and equipment failures can also be introduced in simulation scenarios with minimal risk. As training is extended to new applications, there is a need to define performance measures for tasks which are not practiced in current training programs, including maneuvering in waves, picking up persons in the water (PIW's), and dealing with hazards such as fires. While coxswains are required to be competent in these tasks, there had been no way to assess skills in operational scenarios prior to using simulators. Performance rubrics had to be established to define expected performance in conditions which could now be practiced in a simulator. Simulators have been widely used to assess performance in operational conditions using scenario-based training exercises. Both high and low fidelity simulators have been used to investigate human performance in marine operations (Sellberg, 2017). This study provides an additional case of how a simulator can be used to measure performance in an exercise that would otherwise be prohibitive due to risks and logistics.

An additional benefit of the simulation is the ability to automate measurements using computer software. Information taken from the virtual environment can provide precise information on the position and speed of the vessel and the timing of events. Recorded data provides a means to track events and determine if tasks have been completed in order and within desired ranges of time and distance. Traditionally, lifeboat training has been performed with a live instructor who assesses performance based on observed behavior of trainees as they practice tasks. In emergency scenarios, students are required to complete tasks which require timely decision making while maintaining control of the vessel as it moves on the ocean waves. Measuring competence can require tracking of several factors which can make evaluation difficult. Stanney et al. (2013) note that high instructor workload due to the amount of information that must be measured and processed can reduce the effectiveness of the training environment. Carroll et. al (2008) suggest automated training tools can be implemented to reduce instructor workload and improve training efficiency by focusing information provided to instructors for assessment and for identifying skills that need improvement. Lifeboat simulators have also been deployed as independent trainers (i.e. without a live instructor) and automatically track student performance and assess competence based on measured behaviors in simulator scenarios. To be an effective trainer, automated simulators are expected to evaluate performance as good as a live instructor.

The study had two goals 1) to assess how automated tracking software developed in simulator software scored participants in comparison to a live instructor, and 2) to evaluate a rubric that was created to measure coxswain performance in conditions which are prohibitive in current training.

METHODOLOGY

An experiment was conducted using participants who had received different types of training providing variability in skill and competence. Participants were evaluated by both the simulator and live instructor in a simulated emergency scenario using a common rubric to assess performance.

The study was comprised of three stages. First, a scoring rubric was established to define expected performance of participants on lifeboat launch and maneuvering tasks in plausible emergency conditions. Once completed, automated tracking features were implemented in simulation software to measure participant performance using the rubric. The last stage involved the execution of the experiment to collect data on participant performance. A scenario was created emulating an emergency launch from an oil and gas platform. Participants performed the simulator exercise and were assessed by a live instructor and the simulator using a common rubric. Details on the stages are provided in the following sections.

Scoring Rubric

A scoring rubric was created to identify measures of task performance for launch and maneuvering tasks in seas states and for dealing with hazards such as fires. The criteria for task completion were established by Subject Matter Experts (SMEs) to reflect a standard of proficiency as identified in recognized training standards, including the Standards of Training, Certification and Watchkeeping for Seafarers, and model lifeboat courses (IMO, 2010). Table 1 provides a list of tasks and objectives used to measure performance. Tasks are divided by type and order of execution in an evaluation scenario involving the launch of a lifeboat in an emergency.

Each of the tasks were quantified using subtask measures to provide an objective measure of performance that could be determined by a live instructor and simulator. Measures taken varied by task type. For launch tasks, tasks were procedural in nature, and performance was measured based on ability to complete four tasks successfully and in order. Clear-Away tasks required knowledge of operation of release equipment and tasks had to be completed in order and timely to minimize risk of harm. As an example, to perform a successfully splashdown and release, the participant needed to open the hook within 10 seconds of being in the water and then apply throttle within 5 seconds of releasing hook. On-water tasks required the user to demonstrate ability make decisions on the best way to perform the tasks and maneuver the boat successfully. As an example, for a PIW pickup, the participant had to approach from a downwind position to reduce chance of being pushed into the PIW by wind and waves. The participant then had to slow the boat to below 0.5 knots of speed, which SME's believed to the fastest speed for a safe recovery, beyond which the task was not possible to complete. The lifeboat had to be held in position for 10 seconds, and the number of attempts to keep the boat stopped for this time was an indicator of vessel control. Contact with the PIW would also be harmful. Compass tasks required the user to maintain a heading without veering off course by more than 30 degrees of target heading, and with few corrections of vessel heading when on course, indicating control of vessel.

As tasks became more complex, the measurement of performance considered more factors. For procedural tasks, such as launching the boat, measures were recorded for each task independently while observing the order of tasks. For more complex tasks, such as a PIW pickup, both the simulator and instructor had to assess multiple criteria simultaneously (e.g. direction, contact, speed, number of attempts). The number of simultaneous measures for each task is indicated in Table 1.

Table 1: Scoring Rubric Categories

	Task Name	Task Objective	Subtask Measures	Number of Simultaneous Measures
Launch	Launch lifeboat	Lower vessel to water without stopping, start engine, order use of sprinkler and air system if informed of fire, ensure lifeboat buoyant when releasing hooks	- operate equipment successfully - perform tasks in order	1
Clear-Away	Splashdown and release	Promptly release hooks using hook handle release and apply throttle	- correct order and operation of hook release and throttle - time taken to release hooks, apply throttle	2
	Contact with platform	Maneuver vessel and do not make contact with platform after release	- no contact with launch platform	1
	Clear platform	Safely leave clear away zone by moving away from rig quickly and avoid hazard.	- direction of travel - time to clear hazard	2
On-Water	Navigate by compass	Maintain a compass heading with minimal veer from target heading and control heading under influence of waves	- course of vessel - heading corrections made during transit	2
	Stop at a mark	Approach a static object accounting for wind and wave direction. Use a speed to allow stopping. Stop close to landmark (2-3 boat lengths) and maintain position	- direction of approach - speed at stop time stopped - contact speed - heading at stop - number of attempts	4
	Recover a PIW	Approach a drifting PIW accounting for wind and waves to minimize chance of contact. Use a speed to allow stopping. Stop close enough to PIW to allow pickup and maintain position in waves		
	Come alongside a vessel	Approach a vessel accounting for wind and wave direction. Use a speed to allow stopping. Stop close to vessel (less than 0.5 meters) and at an angle to allow personnel transfer and maintain position		

Measurement and Scoring

A grading scheme was adopted to provide indicators of failure to complete tasks, acceptable performance, or expert performance on completion of task. For measurements requiring time, proximity, and direction, values were assigned based on the expected performance of an expert completing the task. Thresholds were assigned for acceptable measures of performance (i.e. participant was over 50% of expected speed or time), and for measures that would indicate the task was not completed (i.e. participant was over 100% of target speed or time). For each task, participants were scored a value of 1, 3 or 5, based on the completion of subtask measures identified in Table 1. A value of 1 indicated a failure to complete a needed subtask, and hence the overall task was not achieved. A score of 3 indicated the user had completed the subtasks with an acceptable level of performance, with no critical failures. A 5 indicated that all components of the task had been completed successfully with an expert level proficiency. Table 2 provides a sample of the grading scheme applied for a PIW pickup.

Table 2: Sample Grading Schema

Task	1 Point - Failure	3 Point - Acceptable	5 Points - Expert
Recover a PIW	- Was unable to recover a PIW on first attempt or hold position - Could not reduce speed <1 knot. - Contact made with PIW - Approached from upwind and drifted down to the PIW	- Was able to recover a PIW on first attempt - Stopped within 2.5m of the PIW from the side hatch for a minimum of 10 seconds or more - Slowed to within 0.5 - 1 knot - No contact with PIW - Approached from abeam to the wind - Stopped with the bow pointed off of the wind	- Was able to recover a PIW on first attempt - Stopped within 2.5m of the PIW from the side hatch for a minimum of 10 seconds - Came to a complete stop (speed > 0.5 knots) - No contact with PIW - Approached from downwind - Stopped with the bow pointed towards the wind

The rubric established performance using quantifiable measures including proximity to objects, time to complete task, vessel speed, and heading. To record data, both the instructor and the simulator performed in-situ measurements for all tasks that needed to be completed in the evaluation scenario. Scoring features were added to the simulator software to capture data instantaneously and provide a resulting score for each task. For launch tasks, the simulator identified interaction with vessel equipment based on when actions were completed (e.g. capturing time when hook release handle was pulled using the generated electromechanical signal) and the order of tasks completed. Clear-Away and On-Water tasks used proximity and directional zones to capture information on vessel speed, heading, position, and time as the user maneuvered the lifeboat through the virtual environment and entered or exited zones. The captured data was used as input to a scoring algorithm which generated a score based on the grading schema that had been established.

The instructor could assess performance on Launch and Clear-Away tasks through visual observation of trainee actions (e.g. visually observing when the participant pulled the hook release handle). Display cues were added to the simulator instructor station to provide the live instructor with access to the same information being recorded by the simulator for On-Water tasks. The simulator instructor station was modified to provide a computer display of the scoring zones so the instructor had access to the same proximity cues as the simulator for On-Water tasks. A digital display of vessel position, speed, heading, and proximity to targets was given on a 2D map to provide the instructor with the same real time information used by the simulator for scoring. The instructor recorded scores on a paper template and assigned a scoring value using the same scoring rubric. Figure 1 shows a sample of the scoring zones provided on the instructor station.

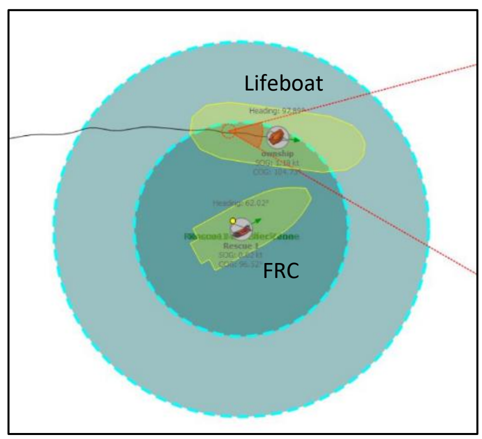


Figure 1: Sample scoring zones showing position of Lifeboat and FRC

Evaluation Scenario and Data Capture

Performance data was collected through a test program using participants who were being evaluated on their ability to perform a successful launch of a lifeboat in an emergency as discussed by Billard et al. 2018b. As the final stage in a year-long test program, participants had to practice launching and maneuvering in an emergency scenario using a simulator as the test environment. This test phase provided the opportunity to take measurements with a live instructor and the simulator using participants with varying levels of training expertise, and using a scenario allowing the rubric to be applied in practice.

As discussed in the study performed by Billard et al. (2018b), participants completed the evaluation scenario until successful on all tasks. Launch, Clear Away and Maneuvering tasks were performed in order. If successful on Launch and Clear Away tasks, the participant could start the scenario with the lifeboat launched to start on-water tasks. Participants performed the scenario until they completed all tasks successfully or until they had tried the scenario a total of six times.

All participants were given the same scenario and were provided with the same scenario briefing. The scenario comprised of an emergency event in weather conditions that were representative of common operating conditions in the North Atlantic. The parameters of the scenario were set to night time with clear visibility, 13 knot winds, and 3 meter wave height. The briefing indicated that an explosion had been heard on the platform, followed by a fire alarm. The Offshore Installation Manager (OIM) had ordered an evacuation from the platform and the duty was to launch the lifeboat and assist in a search and rescue exercise once in the water. Figure 2 provides an overview of the emergency scenario and the tasks that were to be completed. This image was provided to trainees in the briefing. The emergency scenario comprised of launch, clear away, and on-water tasks which were measured by both the simulator and instructor. On-Water tasks included inspecting a Liferaft for casualties, recover of PIW's, and then transfer the PIW's to a Fast Response Craft (FRC). The scenario allowed for practice of all tasks listed in Table 1, in the order listed. On water tasks consisted of two instances of navigation by compass and two PIW pickups. The instructor role-played as an OIM or crew member in specific tasks (e.g. receiving command to turn on sprinkler system) and acknowledged receipt of voice commands. For tasks involving stopping next to an object, the instructor acted as a crew member in the vessel to provide information on proximity to objects, as the vessel operator would in practice rely on voice communication with crew members to gauge distance to objects, such as a PIW, which are difficult to see from the coxswain's viewpoint.

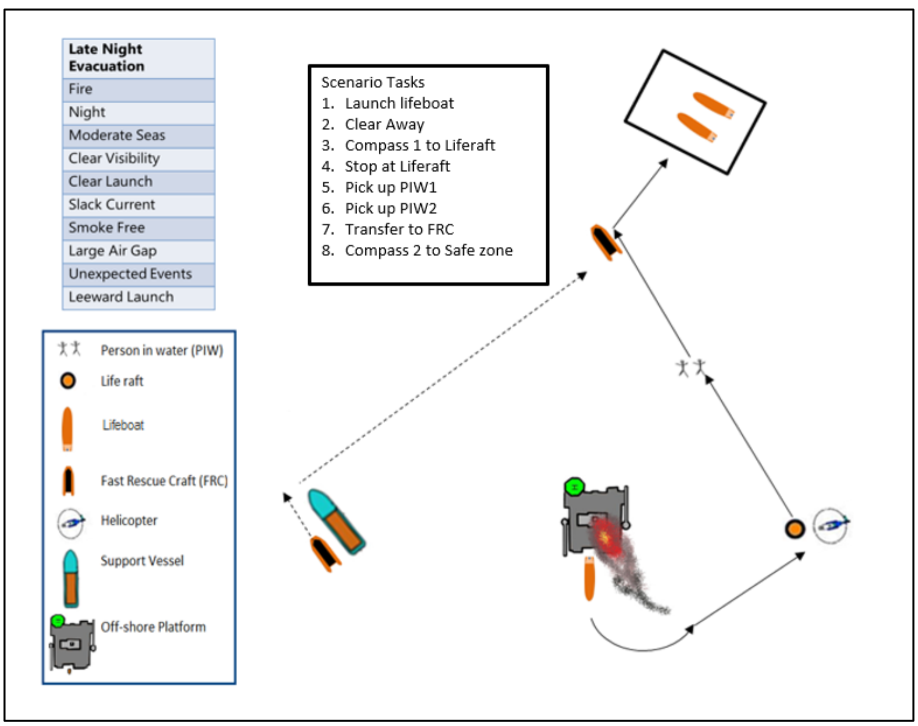


Figure 2: Evaluation Scenario

Simulator

Participants completed tasks in a simulator with a representative layout and equipment of the real lifeboat. The simulator is certified by Det Norske Veritas Germanischer Lloyd (DNV-GL) and Transport Canada as a simulator capable of representing realistic situations needed for training. The simulator provides a visual mockup of the lifeboat and is equipped with real lifeboat equipment (e.g. steering wheel, throttle, brake release, compass) allowing participants to operate the controls needed to launch the lifeboat in a simulation environment complete with visuals and sounds. The simulator motion model, equipment, and layout was modelled to be the same as the real lifeboat. This simulator had been used in research studies to measure skill transfer associated with training to maneuver a vessel (Magee et al., 2016) and skill retention (Billard et al., 2018a, Billard et al. 2018b). The lifeboat modelled in the study is currently used on offshore platforms in the North Atlantic. The lifeboat is approximately 9.4 m long, 3.5 m wide and 6 m high, with a draft of 2.9 m. Figure 3 shows the simulator layout and the lifeboat that was modelled.

Figure 3: Lifeboat Simulator Interior and Lifeboat

Participants and Instructor

Novice participants with no previous experience were recruited for the study. A training program was developed to emulate industry practice of receiving initial lifeboat training at a shore-based facility followed by quarterly practice events. Participants received quarterly training one of three ways emulating current industry practices, including training using live drills, Computer Based Training, and simulation training programs. Three quarterly practice events were performed, followed by an assessment exercise to measure whether skills acquired in the training program transferred to a plausible emergency event. The training resulted in variable performance based on the type of training received. Details on the training provided and the resulting variability in performance are discussed in Billard et al., 2018b.

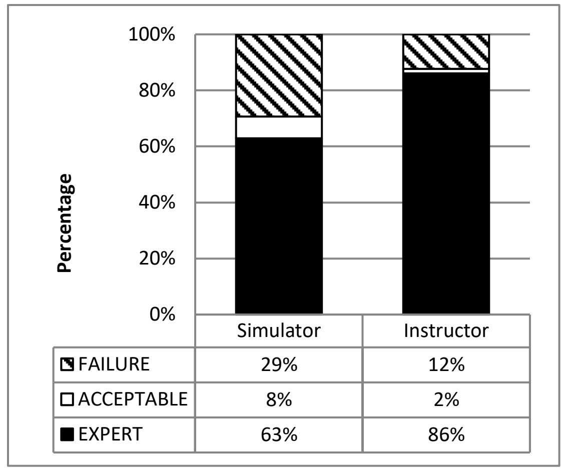
The instructor used in the study was a Subject Matter Expert with experience in small boat handling and personnel training. The instructor was trained on the operation of the lifeboat and simulator prior to conducting the study and was provided time to practice scoring with the rubric prior to starting data collection. On completion of the study, debriefing sessions were held with the instructor to gain insight on how the instructor scored participants based on perceived actions and measures of performance using the rubric. This feedback was used to further investigate differences in scoring between the simulator and instructor and to assess instructor workload.

RESULTS

A total of forty participants completed the evaluation scenario. As each participant was allowed multiple attempts at the scenario, there were multiple scores taken for each individual. In total, data was collected for 81 Launch and Clear-Away attempts, and 124 On-Water exercise attempts. Scoring results are presented as percentages of failures (1) or successes (3-Acceptable or 5-Expert) for all recorded attempts. Scoring results are analyzed by the three tasks types identified in Table 1.

Launch Lifeboat Tasks

The results indicated the simulator recorded more failures than the instructor on procedural lifeboat launch tasks. The tasks were completed in order, and both the simulator and instructor were able to monitor each of the four tasks one at a time. Investigation of the subtask measures indicated the primary difference in scoring was due to the scoring of water entry, with the simulator scoring a failure for over half of the participants, and the instructor scoring a failure in less than 10% of the participants, as shown in Figure 4. The scoring difference was attributed to a combination of the instructor's inability to visually recognize when the lifeboat was in the water and sensitivity of the simulator in measuring when the lifeboat was buoyant. The simulator and instructor had common scores for lowering, engine start, and starting of air and sprinkler system.



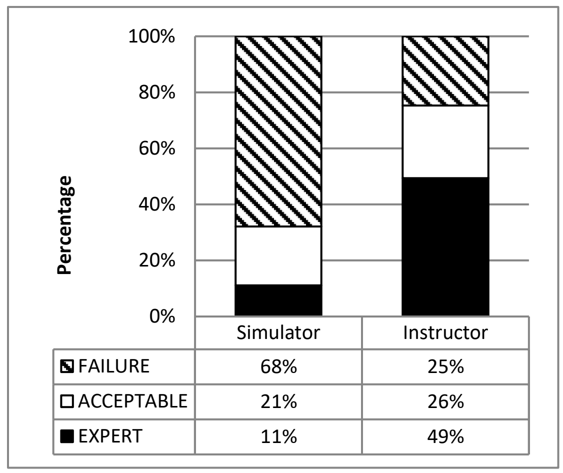
	Simulator	Instructor
FAILURE	29%	12%
ACCEPTABLE	8%	2%
EXPERT	63%	86%

Figure 4: Launch Task Scores

Clear-Away Tasks

The simulator again scored more failures on Clear-Away tasks.

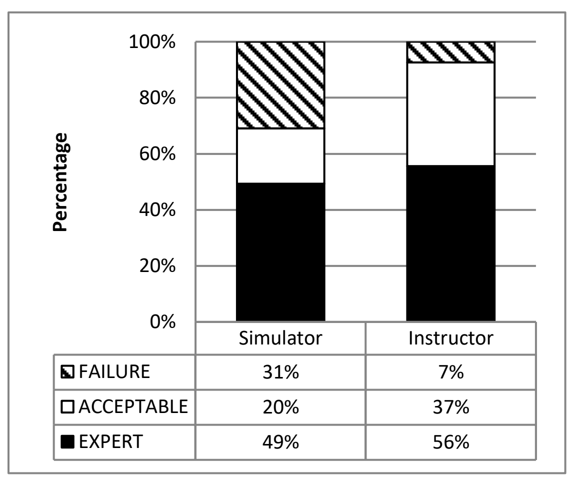
As indicated in Figure 5, on vessel Splashdown, the simulator scored failures over twice as often as the instructor and scored 31% of records as failures on the Clear-Platform task compared to 7% recorded failures by the instructor. Analysis of the recorded subtask measures indicated that the instructor was more lenient on measures which required the participant to complete tasks within a certain time. For example, the instructor gave a passing score (3-Acceptable or 5-Expert) to 23% of participants for time taken to release the hooks whereas the simulator scored a failure (1) for the same participants' performance. The instructor also gave a passing score to 5% of participants for time taken to apply throttle when the simulator scored a failure. For all splashdown subtask measures, the simulator and instructor scored the same for only 37% of participants, with the simulator and instructor scoring a common failure for 25% all participants.



	Simulator	Instructor
FAILURE	68%	25%
ACCEPTABLE	21%	26%
EXPERT	11%	49%

Figure 5: Splashdown Scores

For Clear-Platform tasks, the simulator scored a failure to 21% of participants when the instructor gave a passing score due to measured time to leave the clear away zone, as shown in Figure 6. The simulator also scored a failure for 4% due to the direction taken by the participant when leaving the zone when the instructor scored a pass. The instructor noted that when scoring the primary focus was visual observation of correct operation of equipment, with timing of tasks taken as a secondary measure.



	Simulator	Instructor
FAILURE	31%	7%
ACCEPTABLE	20%	37%
EXPERT	49%	56%

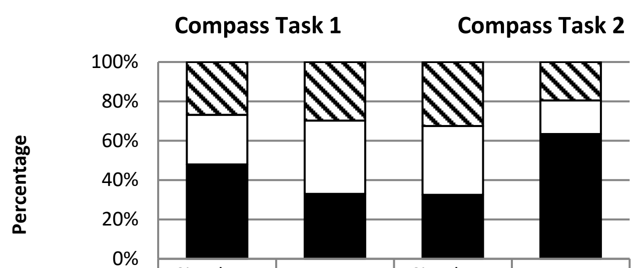
Figure 6: Clear-Platform Scores

The simulator and instructor scored the same for contact made with platform.

On-Water Tasks

The on-water tasks comprised of two types of activities including navigating by compass, which was performed twice, and stopping the vessel near an object, which was performed four times (Stop at Liferaft, PIW1, PIW2, and Stop next to FRC for personnel transfer).

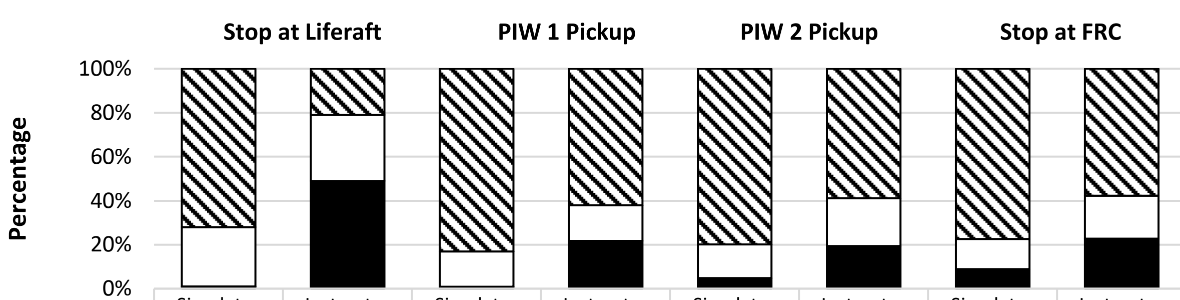
Figure 7 shows the percentages of participants' scores generated by the simulator and the instructor for the compass tasks. The scoring indicates the instructor and simulator scored similarly for the compass runs. Differences in scoring were mainly due to how the simulator and instructor gauged the amount of steering corrections the participants had to make to maintain a desired heading, with the simulator having less tolerance for a high number of corrections. The instructor discounted this behavior as it was felt corrections were needed to deal with instantaneous wave impacts on the vessel. The simulator and instructor scored consistently on the amount of veer and overall ability to complete the task.



	Simulator	Instructor	Simulator	Instructor
FAILURE	27%	30%	33%	20%
ACCEPTABLE	25%	37%	35%	17%
EXPERT	48%	33%	33%	63%

Figure 7: Navigate by Compass Scores

Figure 8 depicts the percentage of participants that received a fail, acceptable or expert performance score for the four stopping tasks. The instructor was more lenient on scoring of tasks involving stopping the lifeboat. The simulator scored a failure on stopping tasks over 72% on each of the stopping tasks. The highest failure rate provided by the instructor was 62% for each of the 4 stopping tasks. The instructor recorded a failure rate as low as 21% for the task of stopping next to a Liferaft. This task had the highest allowable contact and stopping speed in the rubric.



	Simulator	Instructor	Simulator	Instructor	Simulator	Instructor	Simulator	Instructor
FAILURE	72%	21%	83%	62%	80%	59%	77%	58%
ACCEPTABLE	27%	30%	16%	16%	15%	22%	14%	20%
EXPERT	1%	49%	1%	22%	5%	19%	9%	23%

Figure 8: On-Water Task Involving Vessel Stopping Scores

An analysis of failure reasons provided for each of the stopping tasks is shown in Figure 9. These results indicate the primary reason for failures as recorded by the simulator was due to the measured speed of lifeboat on stopping. The simulator also scored more failures due to speed of contact. The simulator and instructor scored consistently on the number of attempts needed to complete the task. Feedback provided by the instructor indicated that it was difficult to perform all measures during stopping tasks as some attention was needed to perform simultaneous performance measures and instructor role-playing (e.g. to act as a crew member and provide distance measures to the coxswain when executing the tasks).

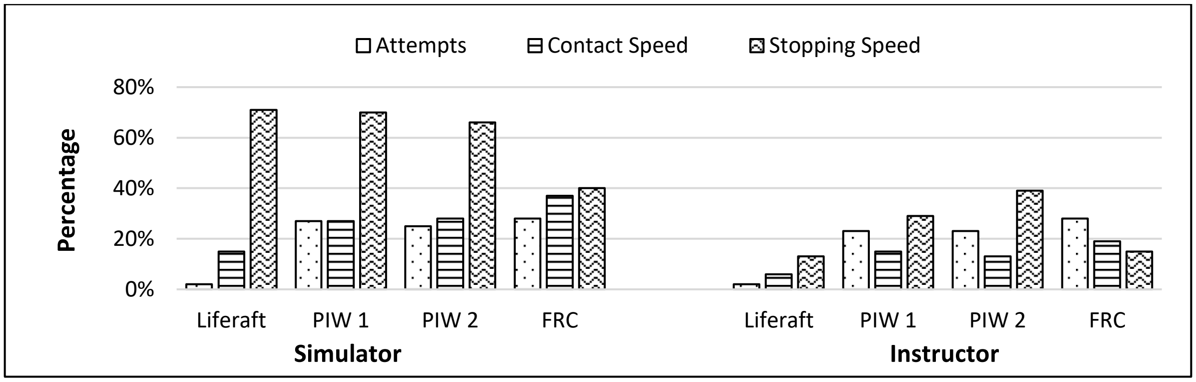


Figure 9: Recorded Failure Reasons for Vessel Stopping Tasks – Percentage of Participants

Discussion of Results

The results of the study allow for assessment of how the simulator scored performance compared to a live instructor and how effectively the scoring rubric could be applied to tasks which have not traditionally been measured.

The simulator and instructor appeared to score performance consistently for tasks involving 1-2 simultaneous measures. The highest difference was observed when 2 or 4 measures had to be taken simultaneously, with further separation when the instructor had to perform role-playing tasks. In debriefing sessions, the instructor noted that it was difficult to track measures simultaneously, including contact speed, and time to hold position, while interacting with the participant in the simulator. The instructor also had difficulty in observing participant behavior on Launch and Clear-Away tasks to measure their ability to use lifeboat equipment, such as the hook release and throttle, while monitoring the time to complete tasks. This feedback and the analysis of failure reasons suggests the workload for the instructor for these types of tasks was high enough to hinder the ability to accurately measure performance.

Further analysis of the On-water tasks measures provided insight on the observed differences between the instructor scoring and simulator scoring. The instructor relied on a combination of data from the simulator and subjective perception of speed to determine if the user had stopped the lifeboat. The motion of the lifeboat in the waves made the perception of speed more difficult, and the instructor relied on subjective visual perception of relative speed to the object by viewing the motion of the vessel in the simulator visuals. The calculated speed of the vessel, as used by the simulator, was modelled to take into account the instantaneous speed of the vessel as it was influenced by wave induced motions and vessel rotation, in addition to propulsion, which can increase or decrease speed at the time of measure. A suggested outcome of the study is to simplify the calculation of speed used by the simulator to provide a more consistent measure as the instructor. This introduces a caveat that the actual speed may exceed thresholds of human performance.

The outcomes the study will be used to further refine performance rubrics applied to operators as they practice for emergency events. For tasks conducted one at a time and in order, the simulator and the instructor score similar and can consistently discern failures using tracked or observed data. The results indicate that measures requiring tracking of simultaneous measures involving position and time are difficult for the live instructor to execute. In tasks involving simultaneous measurements, the simulator has the advantage of being able to track more information. A suggestion is to use the automated tracking to aid the instructor in the evaluation through the provision of tools to provide feedback on time of completion of task and measures of speed and position. As noted by Carroll et al. (2008), this approach allows the instructor to focus on training as opposed to measuring outcomes. An alternative is to prioritize the measures to one or two measures, although this could result in a relaxed measure of competence. The study also provided insight on how an instructor would use the rubric and the simulator to assess performance. Further refinement of the automated tracking tools will improve the scoring accuracy of the simulator and increase the capabilities of the simulator as an automated trainer. Rubrics can also be refined and evaluated in the simulator as more knowledge is gained on human performance in adverse weather and emergency lifeboat launches.

The study indicates the simulator can be used to provide an effective means of performance measurement in emergency situations. This is a novel application of simulation-based training to extend industry's ability to assess their crew competence and readiness. The performance of individuals in moderate weather conditions is of particular interest, as both the simulator and instructor predict a low success rate on slow speed maneuvering tasks in wave conditions which are common in operation. Further studies are planned to extend the training applications to additional tasks, hazards, and environmental conditions and to measure student learning and performance using a simulator as the test environment. Studies will also be performed to assess how automated tracking tools can be used to reduce the workload for instructors, and how the simulator can be used to provided automated evaluations of students to measure skill acquisition and adapt training materials to individual students. As automated simulators are adopted for use in Oil and Gas training, there will be a continued focus on improving the automated tracking tools and scoring rubrics to provide measures of performance that can be used to assess competence of crew and preparedness for emergency situations.

ACKNOWLEDGEMENTS

We thank Petroleum Research Newfoundland and Labrador and the Industrial Research Assistance Program of the National Research Council as sponsors of the study.

REFERENCES

Billard, R., Smith, J., Magee, L., Veitch, B., (2018a), Simulator Training for Offshore Oil and Gas Emergency Preparedness, The international forum for the military simulation, training and education community (ITEC) , Stuttgart, May 2018.

Billard, R. Smith, J., Veitch, B. (2018b) - Assessing Lifeboat Coxswain Training Alternatives Using Simulation, Journal of Navigation - Submitted

Carrol, M., Champney, R.K., Milham, L.M., Jones, D.L., Change, D. and Martin, G. (2008). Development of an instructor aid to diagnose performance. Proceedings of the Interservice/Industry Training Simulation and Education Conference (I/ITSEC), Orlando, FL, 1-4 December 2008.

C-Core (2015) Metocean Climate Study Offshore Newfoundland and Labrador, Nalcor Energy Report. http://exploration.nalcorenergy.com/wp-content/uploads/2016/09/Nalcor-Metocean-Study-Final-Report-Volume-2-27-May-2015.pdf

International Maritime Organization (2010). Revised STCW convention and code adopted at the Manila conference, http://www.imo.org.

Magee, L.E., Smith, J.J.E., Billard, R., & Patterson, A. (2016). Simulator training for lifeboat maneuvers. In proceedings of the Inteservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2016. Paper number 16030.

Sellberg, C., (2017) Simulators in bridge operations training and assessment: a systematic review and qualitative synthes WMU Journal of Marititme Affairs (2017) 16: 247.

Stanney K., Carroll, Champey, R., M.m Devore, L, Kelly, H. (2013). Virtual Environment Training Design, Pros, Cons, and Tailored Solutions. In C. Best, G. Galanis, J. Kerry & R. Sottilare (Eds.), Fundamental Issues in Defense Training and Simulation. Ashgate Publishing Limited: Surrey, England. 135-154.