Heavy Weather Ship Handling Simulation With and Without Motion

Jelte E. Bos,

TNO Perceptual and Cognitive Systems, Soesterberg, Netherlands, and
Research Institute MOVE, faculty of Human Movement Sciences, VU University, Amsterdam, Netherlands, Jelte.Bos@tno.nl

Pierre J.L. Valk

TNO Training and Performance Innovations, Soesterberg, Netherlands, Pierre.Valk@tno.nl

Wietse D. Ledegang

TNO Perceptual and Cognitive Systems, Soesterberg, Netherlands, Wietse.Ledegang@tno.nl

Bart Nienhuis

Defence Materiel Organisation, The Hague, Netherlands, B.Nienhuis@mindef.nl

ABSTRACT

Despite the fact that high risk sailing operations are often caused by heavy weather and characterized by vigorous ship motions, the vast majority of (full mission) bridge simulators are Earth fixed. We therefore evaluated the value of adding motion to ship simulation in terms of technical feasibility, measurability of performance, and effects on training. Of eighteen navy personnel, nine received a 30 min training session with and nine without physical motion in a bridge simulator. Pre and post tests showed no objective effect on training performance. Subjectively, however, physical motion was judged beneficial.

KEYWORDS

Heavy weather; ship handling; simulation; motion; human performance; operator performance.

INTRODUCTION

To train crew for certain tasks, simulation is generally considered a cost-effective alternative to training on the job. Moreover, simulation offers the possibility for applying certain scenarios in well controlled conditions, possibly repeatedly, which control is generally lacking in real life, e.g., at sea where wind and waves generally follow their own course. Apart from cost-effectiveness, simulation also offers the possibility to train high risk operations, which for ethical and safety reasons cannot be trained in real life operations. Interfering with the operator, either for training or to measure performance or stress, for example, is also much easier in a simulated environment than in real life operational conditions. In addition to training, simulation can be of value with respect to the development of (safety) criteria and/or heavy weather handling doctrines. Although incident analyses are indispensable in this respect, simulation offers the possibility to focus on certain aspects. This especially holds for human factors and even more so for high risk issues again being unethical if not impossible to study in real life.

At sea, motion does affect performance in several ways. Impaired posture, gait and manual activities, fatigue and seasickness are just some effects to mention. The vast majority of bridge simulators, however, is fixed based, and literature further elaborating on the pros and cons for moving base sailing simulation seems to be missing. This contrasts with top level (D) flight simulators which are equipped with a motion platform as required by regulations. The difference may yet be surprising for a number of reasons. At sea, for example, the average voyage lasts longer than the average flight, why especially (motion induced) fatigue seems to be a more serious issue at sea than in the air. Heavy weather conditions are also harder to avoid at sea than in the air due to a lack of the third dimension for escape. The omission of physical motion in simulation of heavy weather ship handling conditions is therefore especially surprising.

From a technological point of view, the discrepancy between sailing simulation on the one hand, and flight and driving simulation on the other can yet be understood. In flight simulation, air turbulence can be modeled as a simple additive stochastic error signal and in driving simulation road surfaces are generally not affected by car motion. In sailing simulation, however, each wave has to be taken into account explicitly with respect to the visualisation, to the effect thereof on the (motion of the) ship, and vice versa with respect to the effect of the ship cutting through and thus changing the waves. Making sailing simulation realistic is therefore a technologically more challenging task than driving or flight simulation. Although different hydrodynamic computational codes do already allow for reliable calculation of the ship motions, most of these codes are too time consuming to be realized with a man-in-the-loop in real time at this moment. Yet, some models for larger ships and linearized wave models are already at hand, while in the near future also codes for smaller (planing) craft with nonlinear wave dynamics may become available.

Furthermore we assume that not all kinds of training will require a realistic (motion) environment. Acquiring knowledge and training procedures, for example (Rasmussen, 1983), are likely less affected by motion than training skills. However, there will also be scenarios for training specific skills not benefitting from physical motion, such as manoeuvring a tanker in calm waters. Especially in heavy weather ship handling conditions, however, ship motion affects human behaviour considerably.

Building on the work described by Ten Hove and Roza (2010), we further elaborated on the specific effects of physical motion on training by simulation, focusing on a medium sized displacement ship, for which suitable ship motion code was available. To that end, we also elaborated on a performance metrics to quantify the effects at issue. A secondary objective was to differentiate between novices (cadets) and experienced crew (officers of the watch) with respect to the effect of simulator motion during training.

METHODS

Ship and motion code

For the experiment, a 51m mine hunter was selected to simulate, based on a trade-off between its motility and predictability of its motions by means of a hydrodynamic computer code. Regarding the latter we used the ship motion code FREDYN, simulating the dynamic behaviour of a steered ship subjected to waves and wind. The development of FREDYN is one of the on-going tasks of the Cooperative Research Navies group, of which DMO is a permanent partner (CRNAV, see also www.marin.nl/ web/show/id=70052).

Simulator

The experiment was performed in TNO's Desdemona research simulator (Fig. 1 top) with six degrees-of-freedom of motion, nested in a different way as compared to a Stewart platform. The simulator cabin is fully gimballed and can rotate infinitely about all axes, while it can move vertically along a 2m heave axis and horizontally along a 8m linear arm. The linear arm itself can rotate about its central yaw-axis to generate centripetal forces if needed. More information on this simulator can be found at www.desdemona.eu.

Figure 1. TNO's Desdemona research simulator (top) and generic ship cabin inside view (bottom).

Subjects were seated in a generic ship cockpit (see Fig. 1 bottom), with radar panel, directional control and throttle quadrant.

Outside visuals were rendered by a PC-based image system using five projectors (resolution 1024 x 768 pixels) and five flat screens viewed by the subjects from approximately 1.5m (thus creating an out-of-the-window field-of-view of about 200° horizontally and 32° vertically). Image edge blending and distortion was also computed in the image system.

To squeeze the motion of the mine hunter (as calculated by FREDYN) into the motion envelope of the simulator, a motion filter was used taking the specific position, velocity and acceleration limits into account. Part of this software consists of high-pass filters (also referred to as "wash-out" filters) readjusting the simulator cabin to its initial position after each change of motion or manoeuver so as to allow for optimum freedom for the next manoeuver. These motions should be realized smoothly and preferably below the threshold for perception. As a result, simulator motion generally does not match the real motion in all degrees of freedom, and people may get sick in the simulator while not getting sick in the real environment (differentiating simulator sickness from motion sickness).

To mimic slamming events, transient vibrations were added to the motions as calculated by FREDYN, whenever heave acceleration was above a certain threshold and pitch went down. Slamming events were considered important because they are an incentive for crew to decelerate the ship or alter course.

Motion driving algorithms were implemented in MATLAB Simulink and ship model responses filtered by these algorithms and realized by Desdemona have been successfully validated with navy expert officers beforehand.

Scenarios

Three scenarios were designed with three different purposes; habituation, a test scenario and a training scenario.

The first (*habituation*) scenario lasted for 10 minutes and allowed the subjects to get familiar with the simulator, the ship, its instruments and controls, and the tests performed (most importantly the PDT as described below). In this scenario, wind (19-29 kts) and waves (1-4m) increased slowly over time.

The second (*test*) scenario was used to assess the subject's performance. Here, the ship had to be sailed as shown in Figure 2. This trajectory can be separated into five segments. In the first segment (1) another mine hunter had to be followed at appr. 500 yards for two minutes in head seas. Then course had to be changed from 255° to 015° (segment 2), resulting in a course with stern quartering waves that had to be followed for a little less than 5 minutes (segment 3). Then a 180° turn had to be made (segment 4), after which the subject had to sail "back home" as fast as safely possible (segment 5) for a little less than 3 minutes. The total time spent on this scenario was 10 minutes, and this scenario was applied just before and right after a training scenario.

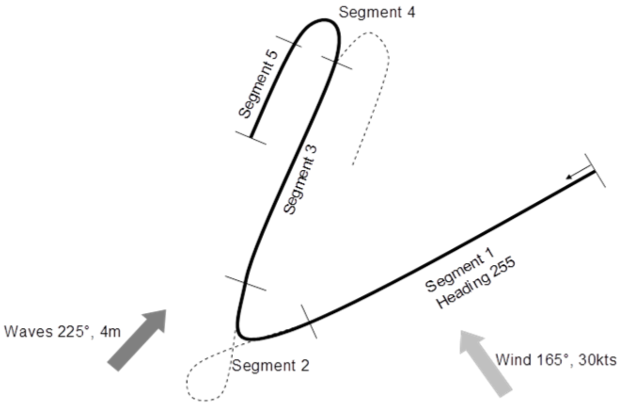

Figure 2. Test scenario with preferred trajectory (solid line) and trajectory as chosen by some subjects instead (dashed lines).

The third (*training*) scenario, lasted for 30 minutes, and consisted of sailing trajectories and turns also included in the test scenario, where the advantages and disadvantages of different options to do so were explained, typically including the risks of slamming, heavy rolling, broaching and surfing.

The habituation and training scenarios were guided by an expert naval commanding officer, allowing ample interaction (by head phones). The test scenarios were guided by an assessor who only gave instructions, while rating the trainee's performance (see below).

Subjects

The study population consisted of a group of 18 healthy subjects (3 female, 15 male) divided into two groups in terms of experience: nine experienced officers of the watch and nine inexperienced cadets. Participation was on a voluntary basis and subjects were not paid. Before the start of the study, participants were fully informed about the objectives of the study, were medically checked (with a focus on neuro-vestibular disorders) and signed an informed consent. The experiment was approved by the local ethics committee.

Measurements

Four different types of measurements were taken: subject ratings, assessor ratings, a peripheral detection task, and measures taken from the simulator loggings.

For the *subject ratings*, trainees were asked to fill out three questionnaires: a "before study questionnaire", an "after session questionnaire", and an "after study questionnaire". The "before study questionnaire" consisted of subjects' characteristics and professional experience. The "after session questionnaire" was composed of different rating scales addressing performance, immersion and situational awareness on a 5-point rating scale (1 = "totally disagree" to 5 = "fully agree", mental effort on a 0-150 continuous scale (RSME 0 = "absolutely no effort", 110 = "extreme effort" (Zijlstra, 1993), and misery on an 11-point rating scale (0 = "no problems at all", 10 = "vomiting" (Bos et al., 2005), all as experienced during the trial. A misery rating of 7="fairly nauseated" or worse was used as a criterion to stop the trial. The "after study questionnaire" assessed the level of realism of the simulation including aspects such as added value for training, operational safety, and performance improvement.

For the *assessor ratings*, a merchant shipping expert, experienced in judging bridge operator performance in a simulator, rated the subjects' overall performance and situational awareness. These ratings were taken for the pre-test and post-test, and, importantly, the assessor was blinded for the training intervention, i.e., did not know whether or not the training was realized without or with motion.

To assess subjects' workload and spare capacity, a so called *peripheral detection task* (PDT) was applied (Van der Horst and Martens, 2010). To that end, subjects wore a headband with a red LED, presenting short flashes in the visual periphery randomized over time with an average interval of 5s (see Figure 3). The subjects' task was to press as quick as possible on a foot pedal. Response time and missed stimuli were taken as measures of workload.

Figure 3. Peripheral detection task (PDT) for measuring workload.

The loggings taken from the simulator during the trials allowed for a fourth set of data, e.g., course, heading, rudder, and 6 DoF ship accelerations, velocities and positions. Based on extensive discussions with expert naval officers a number of measures indicative for their handling qualities were taken from these loggings, such as number of slamming events, experiencing following waves (±15^o) or beam seas (±30^o), and roll angles larger than 20^o.

Experimental design

Subjects were, depending on their professional experience, equally distributed over two experimental groups (F-group and M-group) as presented in Table 1. Both groups followed the same experimental protocol having a habituation phase, a pre-test, a training phase, and a post-test. The only difference for the groups was the training phase. During the training of the F-group the simulator did not move (fixed) and subjects had to perform only having visual input. The M-group underwent their training while the simulator was moving and had both visual and proprioceptive input. Subjects participated in couples per half a day and were blinded for their experimental (training) condition, i.e. they were not told whether their training was with or without motion.

After the briefing and medical check, the subjects went through the scenarios (habituation, pre-test, training and post-test) in turns, the other filling out the questionnaires as applicable to their previous experience.

Table 1. Experimental design. M = with motion, F = fixed , i.e., without motion.

	Habituat.	Pre-test	Training	Post-test
Minutes	10	10	30	10
F-group	M	M	F	M
M-group	M	M	M	M

Data analyses

Repeated measures ANOVA's were applied to analyze the different subject(ive) ratings from the questionnaires, and objective measures from the loggings and PDT, ignoring the habituation runs. Note that the subject ratings are available for all three remaining trials (pre-test, training and post-test), while objective ratings and assessor ratings are available only for the test trials.

RESULTS

General

Three of the 18 subjects (2 cadets and 1 officer) had to stop the trial because of symptoms of misery, and were therefore excluded from the data analysis.

Subject ratings

Cadets showed significantly different learning profiles as compared to officers with respect to their performance (p<.05): cadets showing increased learning while officers stayed more or less stable. These effects were most prominent for ratings on overall performance as obtained from the questionnaires, correct turning of the vessel (see Figure 4) and choosing the right speed. No significant differences were found on pre- and post-test ratings between the two experimental groups (motion vs. fixed base training).

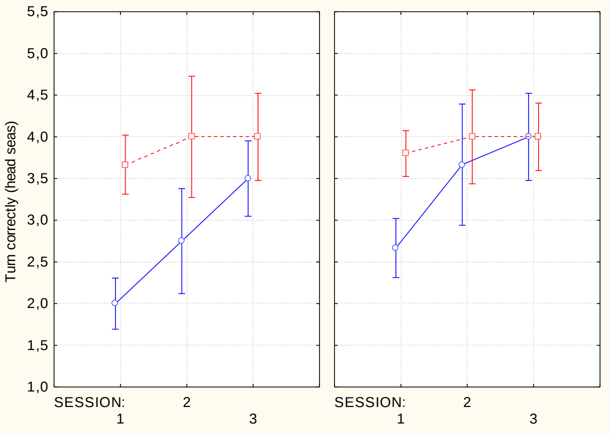

Figure 4. Example result of an averaged subject rating (here the ability to turn correctly in head seas), separated out for session (1=pre-test, 2=training, and 3=post-test), training with motion fixed (F, left) and with motion (M, right), and experience (C=cadets, blue lines and O=officers, red lines). Error bars indicate SEM.

With respect to immersion during the training sessions, subjects in the motion group had significantly higher positive ratings as compared to subjects in the fixed base group ($p<.05$). These effects were most prominent for the level of realism (see Figure 5), the use of motion felt, and the match between visual and perceived motion.

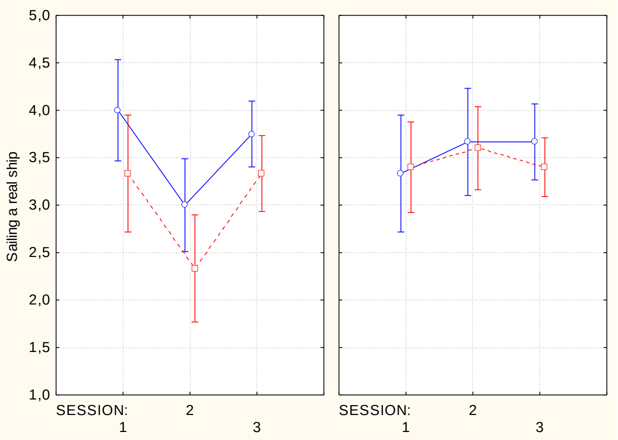

Figure 5. Averaged data for the experienced realism presented analogous to Figure 4.

No significant differences or interaction effects were found on situation awareness between the two experimental groups (motion vs. fixed).

Mental effort scores on the rating scale mental effort (RSME) during training in the motion training group were significantly higher ($p<.05$) as compared to fixed-base training.

After completion of all runs, subjects were asked to complete a questionnaire, giving the following results:

- 75% of the subjects judged the scenarios as being realistic (score 4 or 5 on a 1-5 point scale).
- 75% of the subjects considered that adding physical motion to ship simulators is of added value (score 4 or 5 on a 1-5 point scale).
- 60% felt that during the experimental trials the safety of the ship, crew and cargo was, at least in one instant, jeopardized too much (score 4 or 5 on a 1-5 point scale).
- About 60% of the subjects agreed with the positive advantageous effects of heavy weather handling training (score 4 or 5 on 1-5 point scales for multiple questions addressing this issue).
- All subjects agreed with the statement that due to the decreasing experience in general (decreasing number of ships, smaller crews, crew members spending less time at sea, avoiding heavy weather whenever possible due to improved forecasting methods/tools, etc.) on heavy weather handling, a ship type dedicated heavy weather handling doctrine would be beneficial.

Assessor ratings

No differences were found between training interventions (post- vs. pre-tests). Also no differences were found between cadets and officers.

Logging measures

Figure 6 shows a typical example of a track with performance parameters. Figure 7 shows the PDT results for the same track and subject. No significant differences between the Motion and Fixed groups, nor between cadets and officers were found for any of the parameters as analysed.

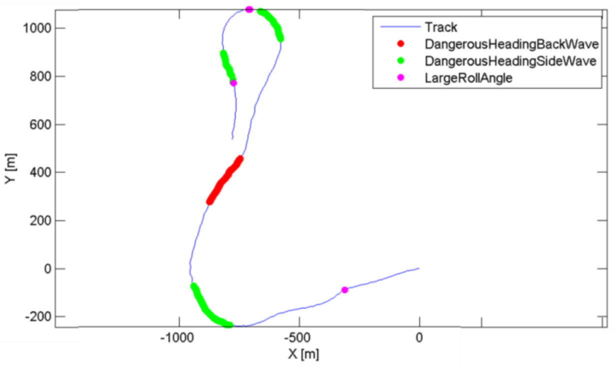

Figure 6. Typical track plot with three performance measures highlighted.

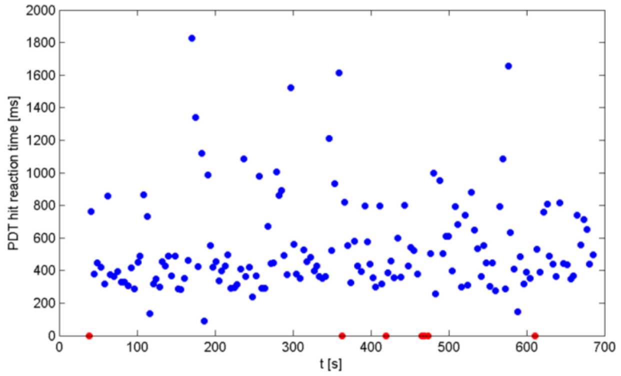

Figure 7. PDT results for the subject sailing the track shown in Fig. 6. Blue dots show the response times and the red dots the missed flashes.

DISCUSSION

The main objective of this study was to demonstrate the feasibility of adding realistic physical motion to a bridge simulation. We focused on a medium sized naval ship (a 51m mine hunter) for which a suitable (hydro-dynamic) ship motion code was available. The results were to be disseminated not only by means of this report, but also by organizing a demonstration for defence stakeholders who could actually experience a trial themselves. Furthermore we aimed at the question whether a distinction could be made between training in a simulator without and with physical motion, making use of both subjective and objective measures. A secondary objective was to differentiate between novices (cadets) and experienced crew (officers of the watch) with respect to the effect of motion during training. To discriminate between the differences at issue, a major effort in this study concerned the development of a performance metrics consisting of subjective self and assessor ratings as well as objective parameters derived from the simulator loggings.

75% of all participants considered the scenarios realistic. Also 75% considered physical motion to be of added value for bridge simulation training. We therefore conclude that we succeeded in creating a training environment that sufficiently satisfied the prerequisites to achieve the main objectives of this study.

During the training sessions, the group who trained with physical motion felt considerably more immersed during the simulator training than the group who trained without motion. The assessor ratings, on the other hand, did not differentiate between the motion and no-motion training groups, nor did the objective measures based on the loggings. The discrepancy between subjective and objective ratings may have three causes: insensitivity of the metrics, an inadequate experimental design, and/or a non-existing effect. Here, we feel that, despite ample reasonable arguments to perform the experiment as designed, especially the training duration was too short. Another possible reason may have been the size and composition of the groups. In this case we had a mixed group of subjects with respect to naval experience. This assumption seems to be validated by the observation that cadets did show an improved (albeit subjective) performance over the three trials (pre-test, training, and post-test), while more experienced officers tended to stay on the same level. Yet another reason for the mentioned indifference concerns the task used, which allowed for instrument readings and therefore rule-based behaviour, instead of skill-based behaviour. Here, motion likely affects skill-based behaviour more than it affects rule-based behaviour. A last reason is related to the way we applied our metrics, in particular the assessor ratings of the pre- and post-tests. Although we deliberately selected an assessor experienced in judging operator performance in bridge simulator training, he had no experience sailing a mine hunter, where a naval observer aware of the specific risks associated with sailing this particular ship might have judged differently.

Furthermore we have been able to design an extensive performance metrics. Although this metrics did not discriminate between effects of motion on all dimensions, particularly the objective measures, it does give a sound basis for future studies. We consider this a major step forward in relating training to performance not only in simulator environments, but also at sea.

Apart from the training duration likely having been too short to make a marked distinction between training with and without motion, the current study also revealed a number of other shortcomings. Because slamming was considered an essential cue for safe ship operation, and slamming was implemented in a rather rudimentary way, the effect thereof should be studied in more detail, especially when smaller (planing) craft are at issue. Likewise, shipping of water onto the bow and deck were judged not to be simulated adequately. Although the current conditions were on the edge of surfing and broaching, obvious mishaps did not happen, and this may be included in further training as well. Lastly, three of the subjects were excluded because of sickness, which may have been an issue too. Note that we here assume these subjects actually did suffer from seasickness and not from simulator sickness as defined above, because the simulator motions apparently were close enough to the real motions as became evident from the subjective ratings in this respect.

CONCLUSION

With this study we showed that adding physical motion to heavy weather ship handling simulation is feasible. Moreover, such an environment can be used for training purposes, the added physical motion having been shown favourable over training in a fixed base simulation environment. Further objectifying the effects at issue would be useful, and the current study also presented (part of) a metrics possibly serving this purpose. Apart from the benefits of adding physical motion to a training simulation environment shown in this paper, we assume the addition of physical motion to be evident for the development of (safety) criteria and/or heavy weather handling doctrines.

ACKNOWLEDGEMENTS

The work presented in this paper was performed within the framework of the programme V937, Improved Performance at Motion, commissioned by the Dutch Ministry of Defense. We furthermore acknowledge our supporting experts Michel Lambregts, Johan van den Arend and Peter Kluytenaar, the naval officers and cadets serving as subjects, and Dick ten Hove from the Maritime Research Institute, Netherlands, and Mark Wentink from the Desdemona B.V., who contributed to the implementation, and realization of the experiment.

REFERENCES

Bos JE, MacKinnon SN, Patterson A (2005) Motion sickness symptoms in a ship motion simulator: effects of inside, outside, and no view. Aviation Space and Environmental Medicine 76:1111-1118.

Rasmussen J (1983) Skills, rules, knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE Transactions on Systems, Man and Cybernetics 13:257-266.

Ten Hove D, Roza M (2010) research into the effects of motion simulation on tugboat captain training and performance in bridge simulators. International Conference on Human Performance at Sea (HPAS) Glasgow, Scotland, UK, 16th-18th June 2010.

Van der Horst ARA, Martens MH. (2010) Ch. 4: The peripheral detection task (PDT): on-line measurement of driver cognitive workload and selective attention. In: Performance metrics for assessing driver distraction - The quest for improved road safety.. Rupp GL (ed) SAE International, Warrendale Pennsylvania, USA. :73-89.

Zijlstra FRH (1993). Efficiency in work behaviour. A design approach for modern tools. PhD thesis, Delft University of Technology. Delft, The Netherlands: Delft University Press.