P r o c e e d i n g s
16th IAMU Annual General Assembly
Opatija, Croatia, 2015

Sveučilište u Rijeci
Pomorski fakultet Rijeka
University of Rijeka
Faculty of Maritime
Studies Rijeka

IAMU
International Association of Maritime Universities

RESEARCH USING MARINE BRIDGE SIMULATORS – HOW WELL DO THE RESULTS TRANSLATE INTO REAL WORLD OPERATIONS?

Dunham, Richard*; Lutzhoft, Margareta; Brooks, Ben

Australian Maritime College,
University of Tasmania
Australia

Abstract. Marine Simulation has been used for training seafarers since the 1960's [1], and much has been published on the use of marine simulation for the purpose of training seafarers. The functional requirements for these simulators are established in the Standards of Training Certification and Watchkeeping (STCW) Convention [2]. Further use has been made of the simulators for the purposes of research, and although there are functional requirements for training, no such requirements have been created for research. How is it possible to determine how the research results from a training simulator translate into real world operations when there are no standards for the capabilities of the simulator?

The classification Society DNV-GL AS, has created technical standards to show that a particular simulation facility is capable of providing the mandatory training as required by the STCW convention [3].This standard includes requirements for realism, and is defined in the document as "the degree the simulator looks and feels like real equipment" [3] p 10 The standard also includes an assessment of the Dynamic Behaviour of the system. In essence these standards are designed to ensure that the trainee in the simulator experiences a realistic environment, and uses tools which may be generic in nature, but are used in the same way, and provide the same information as similar tools used at sea. The DNV-GL technical requirements are designed to allow a maritime administration to show that a marine simulator system meets the requirements of the Standards of Training Certification and Watchkeeping (STCW) Convention when used for mandatory training and assessment. Section A-1/12 of the STCW convention, [2] states in section1.2 that the simulator "be capable of simulating the operating capabilities of shipboard equipment concerned, to a level of physical realism appropriate to the training objectives." Are these poorly defined standards of physical realism and dynamic behaviour sufficient to translate research results to real world operations?

This paper will look at the question of the validity of the research results obtained from a bridge simulator, with a view to establishing any differences and gaps between the simulation requirements for training and the requirements for research, especially research which is industry or user driven. It will follow a typical research project carried out in a marine bridge-simulator, and establish the validity [4] and reliability which may be used for quality control of results following the use of simulators for research. It will further question whether this produces a result that provides external validity, allowing the results of the research to be directly related to real world operations.

Key words: simulation, validity, research

***Corresponding author**

Phone: +61(0)3 6324 9764
e-mail: rcdunham@utas.edu.au

1 INTRODUCTION

Marine Simulation has been used for training seafarers since the 1960's [2], with IMO requiring the use of simulation for training with the introduction of the Standards for Training, Certification, and Watchkeeping Convention in 1978 (STCW78). Since then, much has been published on the use of marine simulation for the purpose of training seafarers. The functional requirements for these simulators are now established in the Standards of Training Certification and Watchkeeping (STCW) Convention as amended [2]. These functional requirements, in short, refer to the simulator being suitable for the selected training objectives and training tasks, as well as ensuring sufficient behavioural realism. The requirements are written in general terms and are open to interpretation. For a flag state to approve a particular simulation facility as meeting the training requirements of IMO, the classification society DNV-GL has produced a standard for Maritime Simulator Systems [3] as "a method of carrying out such approval". Section 1.1.2.1. of the DNV-GL standard states "The purpose of the standard is to ensure that the simulations provided by the simulator include an appropriate level of physical realism in accordance with recognised training and assessment objectives". This is further supported by the International Marine Contractors Association (IMCA) who has published "Guidance on the Use of Simulators" [5]. The purpose of this document is to give guidance on simulators limited to the use for training and competence in the marine contracting industry. It states "It should be noted that the level of realism of each simulator will directly impact the effectiveness of the learning experience for the trainee".

If this same "training" simulator is to be used for research, how well do the results match what will be experienced in the real world?

2 RESEARCH USING SIMULATORS

Research using Marine simulators can be divided into two broad categories. The first is port and harbour development, involving the creation of new ports, extensions or adaptions to existing ports, or research looking at manoeuvring larger ships in an existing port. The second is the research looking at human interactions, either as members of a team, or individual's interaction with technology. The research approach for the two areas differs and so will be looked at separately.

2.1 Research for Ports and waterways

Ankudinov et al [6] noted that there is a value in using ship manoeuvring simulators to support harbour and waterway development. Working Group 20 of the Permanent International Association of Navigation Congresses (PIANC) reported in 1992 on The Capability of Ship Manoeuvring Simulation Models for Approach Channels and Fairways in Harbours [7]. This work is being extended in Working Group 171, with the title 'Ship Handling Simulation Dedicated to Channel and Harbour Design' [8]. Working Group 171 is not expected to complete their work until 2018.

In the PIANC report on Ship manoeuvring simulation models [7], the authors note that "attention in ship manoeuvring studies is generally focussed on the validity of the mathematical ship manoeuvring model. Although this is a very important aspect, other aspects deserve also attention. This applies to a proper problem formulation, the experimental design method, the choice of subjects (pilots and/or masters), and the method of data analysis and drawing conclusions form the investigation."(pp7). The report also details the various techniques used in ship manoeuvring simulation, and the uses to which they can be applied (see Figure 1). The report discusses at length the advantages and disadvantages of each method and notes that a ship manoeuvring simulator has advantages in allowing real-time simulation which will then demonstrate the influence of human reaction time on the events simulated, although this will introduce a stochastic element to the results due to the variability in human reaction times. The authors also note that there must be similarity between the simulator outfit and a real ship. In section 9 of the report, the authors' state "Compared to other models (civil engineering)..... ship manoeuvring models have probably a high degree of validity. With all these models it should be realised that, although they are not perfect, they increase our knowledge with regard to a problem very much. Doing nothing means knowing nothing, thus increasing the chance of errors."(pp27)

The International Association of Marine Aids to Navigation and Lighthouse Authorities (IALA), in Guideline 1058 reviews the use of simulation as a tool for waterway design [9], and suggests the use of simulation to supplement the "existing qualitative and quantitative IALA risk assessment tools". The Guideline states that "the purpose of simulation in AtoN planning and waterway design is to test, demonstrate and document various scenarios for deployment of various AtoN and waterway design under different conditions with the aim of identifying optimal operational safety and efficiency". The Guideline further provides information on the selection of different simulation tools for a given task, and the capabilities and limitations of the different simulation tools, and goes on to discuss accuracy and realism considerations.

A further PIANC publication, Approach Channels, A Guide for Design [10], does not discuss the mathemati-

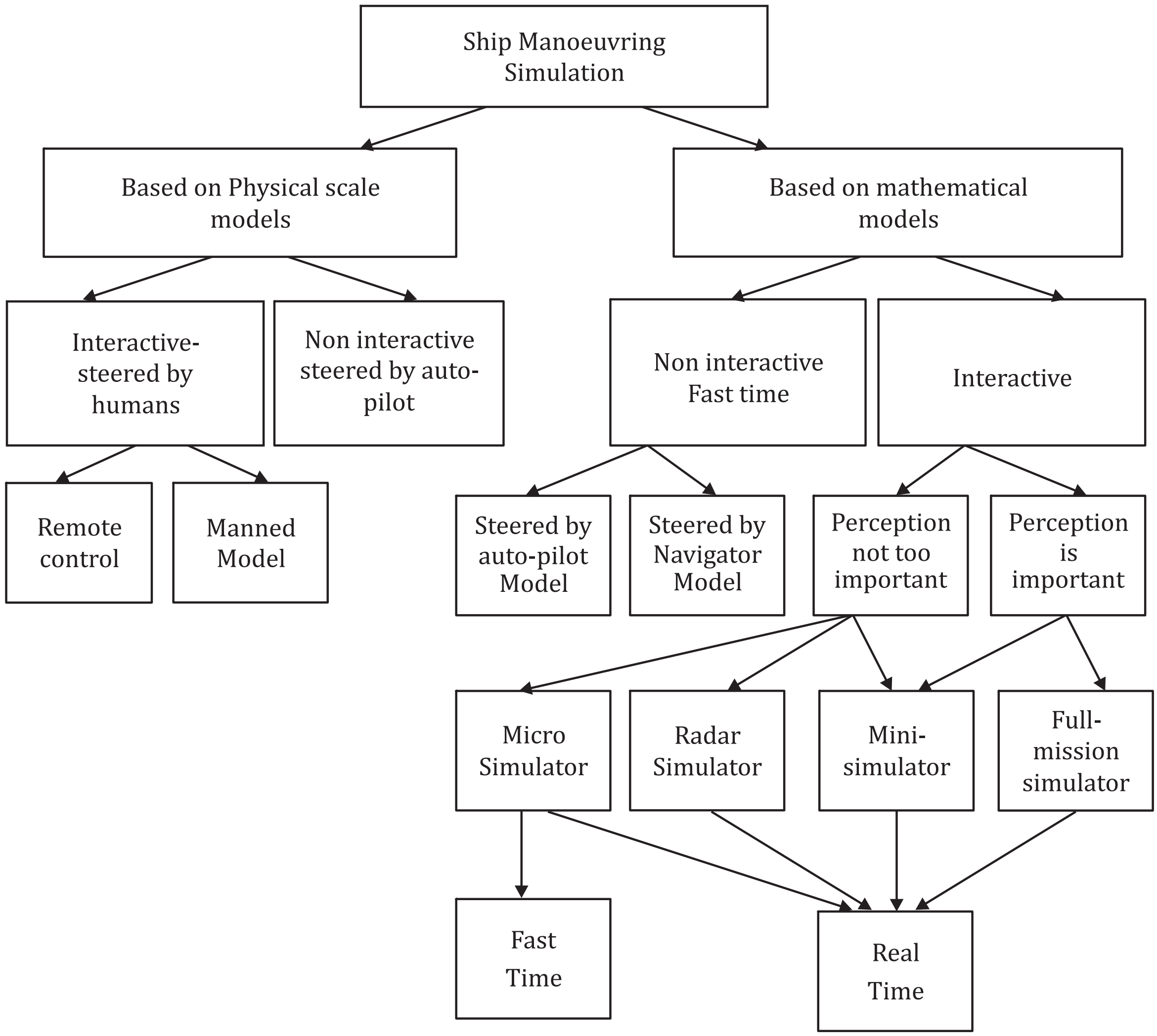

Figure 1 Various techniques used in ship manoeuvring simulation
From Capability of Ship Manoeuvring Simulation Models for approach channel and fairways in Harbours (PIANC)

cal models used, but limits the discussion to the use and value of a Ship Manoeuvring Simulator to a designer. The authors note the importance of the selection of participants, acclimatisation to the simulator, careful recording of performance of the bridge team, and thorough de-briefing following a test run.

2.2 Human interactions

When carrying out a research project looking at human behaviour, perhaps the interactions between humans, their interactions with machines, or their behaviour carrying out a task, then the environment in which this project takes place will have an effect on the outcome. Anderson et al [11] compared the results of laboratory and field studies related to a wide variety of psychological phenomena, and discovered that studies conducted in the laboratory and those in natural settings lead to the same conclusions about human nature: "the psychological laboratory has generally produced psychological truths, rather than trivialities". Research in this area in the Maritime domain could be carried out by observation in the real-world, with the problem that there will be no control over the incidents presented to those involved in the trial. To control this variable, the situation for the trial is switched to a Ship Bridge Simulator, effectively a laboratory. This now defines the location of the research, but does not describe how realistic the environment needs to be.

3 ACCURACY OF THE SHIP BRIDGE SIMULATOR RESULTS

The PIANC publication Approach Channels, A Guide for design [10] refers to "Shiphandling Simulation: Application to Waterway Design" [12]. The authors of this volume state "a simulation will be considered accurate if it can produce piloted track predictions that are useful as a basis for a design decision concerning navigation and risk. Accepted guidelines for this accuracy apparently do not exist, and the accuracy requirement varies depending on the exact nature of the design problem."

None of the standards reviewed here ([2], [3], [5], [7], [9], [10]) give an indication of the measure of the accuracy with which the simulator reflects the real world, but that the accuracy should be checked. Webster et al [12] discuss this as measuring fidelity: the measure to which the simulation matches the "real" situation. The report uses the word fidelity, as it "refers to the appearance and functionality of the simulator as experienced by the pilot". However, measures of fidelity will also include the mathematical model of the ship including the hydrodynamic coefficients used in its definition, the bathymetric model of the port, the visual model of the port, the realism of the bridge etc. The report [12] continues "Ideally, the pilots are provided an environment that so closely resembles a ship's bridge… that they are unable to detect that they are not aboard a ship. In other words, the ideal is a bridge that looks, smells, feels, moves, and sounds like a real bridge, and has views through the windows and ports that are absolutely lifelike. Such an environment would be referred to as having "perfect" fidelity". The environment presented in a ship simulator will not achieve perfect fidelity, in part because there will always be clues that the pilot is not on board a ship, such as lack of movement, incorrect smell, visual scene as a computer generated image and so on. There does not appear to be a method to quantify the fidelity of the situation, other than by questioning the participants, and asking questions as to how well they fell they are on a real ship. This then leads to the stochastic influence in the results depending on the human subject taking part. This then leads to another question for the researcher to answer: Would reliable results be gained from using a larger group and assessing the results statistically improve the reliability of the results?

Kirk and Miller [13] describe the need for qualitative research to be objective, and they partition objectivity into two components, *reliability and validity.* They state "Loosely speaking reliability is the extent to which a measurement procedure yields the same answer however and whenever it is carried out; validity is the extent to which it gives the correct answer"

In their book "Scientific Method – optimising applied research decisions" [14] Ackoff et al state "The reliability of the model can be measured by estimating the variance (or some other appropriate statistic) of the deviations of the observed outcomes from those that were predicted. There are no simple criteria for determining whether the variance is too large: that is, whether or not the model is sufficiently reliable" (p394).There is thus no simple mathematical result from the data produced from the research which will show that it accurately mimics the real-life situation. It would be possible to show that the results have a certain variance, but how does this variance affect our ability to show that the results mimic the real situation.

There are two issues of validity to consider, those of Internal and of external validity. Leedy and Ormrod [4], define internal and external validity as follows. "Internal Validity of a research study is the extent to which its design and the data it yields allow the researcher to draw accurate conclusions about cause-and-effect and other relationships within the data", and "External validity of a research study is the extent to which its results apply to situations beyond the study itself-in other words, the extent to which the conclusions drawn can be *generalised* to other contexts". These two aspects of validity need to be reviewed for the research project to ensure that it is possible to draw meaningful and defensible conclusions. Slack and Draugalis [15] define three steps in establishing internal and external validity. The first step is to assess the statistical conclusion, attempting to show that the results are not due to chance. Only if the results pass the statistics test, should the internal validity be assessed, this time on the basis of the experimental design and operational procedures. The final step is to review the external validity, principally looking at the inclusion and exclusion criteria and characteristics of the study participants. Although the paper is concerned with pharmaceutical experimentation, the process will be similar for research carried out in a simulator, namely that the results produced must be acceptable before internal validity can be reviewed. If the results are accepted for internal validity then an assessment of the external validity can be made.

Osman Balci in his chapter in the Handbook of Simulation, [16], titled Verification, Validation and Testing, states "The question is not to bring *a* solution to the problem, but to bring a sufficiently credible one that will be accepted and used by the decision maker(s)". How then, can the solution be shown to be credible? Perhaps this is the point: if the simulator is designed to reflect the real-life situation, and the models are shown to behave as the real ship will, then the results will translate directly to the real-life scenario. This is an area for further research, and at this stage the solution does not appear to be a trivial.

4 A TYPICAL RESEARCH PROJECT

In his book, Capt Henk Hensen [17] gives a series of 7 steps to setting up a research project, a phase he calls the Validation Phase. This process concerns the confirmation of the accuracy with which all parts of the simulation match the real-world, and uses the term validation with the meaning of a reliability assessment to show the fidelity of the simulation. The final assessment of validity in this case is carried out by professional mariners with experience in the port and with experience in the size and type of ship being used. The author notes "It is important to keep in mind, however, that professional mariners in general only have a feeling – i.e. a subjective indication – of a (simulated) ship's or tug's manoeuvring performance, and cannot provide objective performance criteria. For instance they may feel that speed deceleration at full astern is too fast, but they are seldom able to quantify to what extent". The book further goes into significant detail concerning the methods of confirming the accuracy (validity) of the simulation.

4.1 A Simulator Project

This is an example of how a potential research project used by clients to review a scope of works might be approached by a simulator centre operators.

Initially a client contacts the simulation department and discusses the reason for the project. In this case, the project is to look at the feasibility of extending the berths in a port, to accommodate larger tonnage. The plan also includes some dredging work. The creation of a new bathymetric model was deemed too expensive for the early scoping study, and it was agreed that a depth of 12 metres be applied across the whole model area, a value that could be entered into the operating system at 2 key strokes. This was the designed dredged depth for the port area, and the client was happy that this would not affect the outcome of the trials. The new berths were inserted as simple visual models giving visual clues as to the dimensions of the new port, and these were simply added to the existing port model. The client was offered the possibility of having a ship model accurately modelled for the new tonnage, but decided to use an existing model of similar dimensions. The next discussion was to decide on what the trial was trying to determine. The client, after some discussion wanted to know if the current masters and pilots would be able to manoeuvre larger ships in the port, and to test the environmental limits for the berthing and unberthing. The client provided wind and tide data for the port, and this was selected before each trial was run. The feasibility study could then be run cost effectively to the client's satisfaction.

Following the trial runs in and out of the port by a number of pilots and masters familiar with the port, it was decided that the larger vessel could safely use the new port, with a limit established to the acceptable wind speed. Was this a reasonable assumption?

The question can be answered in terms of the internal and external validity of the research. Was there sufficient data to show that there were no other possible explanations for the results, and can we be certain that the conclusions are warranted by the data [6]? If the answer is yes, then we have satisfied the internal validity of the research. Finally, the external validity: can the results be transferred to the real world, and here there is a problem. With the client deciding that the ship model is not based on the new tonnage to be engaged, that the bathymetric model is inaccurate, that the port visual model has been changed, can the conclusions drawn be transferred to the real-world situation? Referring back to Osman Balci's comment [16], we should ask are the results sufficiently credible for a decision to be made? The answer we leave to the client.

5 CONCLUSIONS

With no standards for Ship Bridge Simulators used for research purposes, we are generally reliant on the simulators produced to fulfil training requirements. The industry requirements for the use of simulators to support a decision making process in port development suggest that the simulator should be as accurate as possible, leaving the final decision on how much reliance to put on the results in the hands of the body initiating the study (the client). It is possible that further research will identify metrics for establishing the validity of a simulation study, or for setting the requirements for validity to be met during the research programme. This may in turn lead, perhaps, to the capability of applying a confidence level for the external validity of the research programme, which would aid decision making.

REFERENCES

[1] Committee on Ship-Bridge Simulation Training, National Research Council, 1996. *Simulated Voyages: Using simulation Technology to train and licence mariners,* Washington DC: National Academy Press.

[2] IMO, 2010. *The International Convention on Standards of training Certification and Watchkeeping for Seafarers as amended*: IMO.

[3] DNV-GL, 2014. *Standard DNVGL-ST0033:2014-08 Marine Simulator Systems,* Hovik, Norway: DNV-GL.

[4] Leedy P, Ormrod J, 2010. *Practical rsearch – planning and design 9th Edition,* Boston: Pearson Education International.

[5] IMCA, Guidance on the use of Simulators, IMCA C 014 Rev 3 Sept 2013.

[6] Ankudinov V, D. L. H. C., 1989. *Use of simulators in harbour and waterway development.* Boston, Ports'89.

[7] PIANC1992, Capability of Ship Manoeuvring Simulation Models for Approach Channels and Fairways in Harbours.

Report of WG No 20 of Permanent Technical Committee II, Supplement to Bulletin No 77.

[8] PIANC, 2014. *Sailing Ahead Oct 2014.* [Online] Available at: http://www.pianc.org/edits/sailingaheadoctober2014a.htm [Accessed 15 March 2015]

[9] IALA, 2011. *Guideline 1058 Use of Simulation as a tool for Waterway Design and Aids to Navigation planning,* Paris: IALA.

[10] PIANC Approach Channels – A guide for design, 1997, Final report of the Joint Working Group PIANC and IAPH, in co-operation with IMPA and IALA. *Supplement to Bulletin No. 95.*

[11] Anderson C, Lindsay J, Bushman J, *Current directions in Psychological Science,*1999, Volume 8, No. 1, p. 3.

[12] Webster WC (editor), Shiphandling Simulation: Application to Waterway Design, Committee on Assessment of Shiphandling Simulation, National Research Council, National Academy Press, Washington D.C., 1992.

[13] Kirk J, Miller M, 1986, *Reliability and Validity in Qualitative Research*, Sage University paper series on qualitative research methods, Vol. 1.

[14] Ackoff R, Gupta S, Minas J, *Scientific Method*, John Wiley and Sons, New York, 1962.

[15] Slack M, Draugalis, Establishing the internal and external validity of experimental studies, *American Journal of Health-System Pharmacists*, 2001, Vol. 58, p2173.

[16] Banks J (ed.), Handbook of Simulation, A Wiley Interscience Publication, New York, 1998.

[17] Hensen Capt H, Ship Bridge Simulators – A Project Handbook, Nautical Institute, London, 1999.