# STCW AND ASSESSMENT OF COMPETENCE BY SIMULATOR: TEN YEARS ON – WHY NO GLOBAL ACCEPTANCE OF THE PRACTICE?

Professor P.M.P.Muirhead Ph.D, FNI (World Maritime University, Sweden)

**Abstract:** More than ten years have passed since amendments to the STCW 1978 Convention introduced new requirements that allowed for the use of simulators in training and assessment. In particular, the Convention made specific provision for mandatory simulator performance standards. As well, STCW Codes prescribed other provisions and requirements for any certificate, that are to be complied with in respect to any 'assessment of competency' and to any demonstration of 'continued proficiency', by means of a simulator.

Despite the passage of time and considerable research by simulator operators in the past decade or so, there is little evidence of any mariners having their competence determined by simulator performance or of having any such outcomes officially recognised by maritime authorities for the issue of a certificate of competency. Why is this so? Are simulators not complying with standards? Is there a lack of approved assessors? Does simulator based competence assessment lack reliability and validity?

The paper examines and evaluates approaches made by a number of simulation researchers over the past decade and asks why the maritime industry has failed to establish a clear argument for simulation technology to be used and accepted for competency outcomes in a similar way to the aircraft industry. The latter readily accepts, on a global basis, the competence of pilots to fly aircraft types after they have undergone flight simulator training and checking for 'continued proficiency'.

Is it realistic to expect marine simulators across the globe to be used uniformly by different assessors, against an agreed set of performance criteria, to measure seafarer competence? Can the competence of individuals to perform functions and tasks safely and effectively by simulator ever be more than achieving satisfactory training outcomes, i.e. experiencing different situations? The paper considers the above questions and draws a number of conclusions.

## 1. STCW, COMPETENCE AND THE USE OF SIMULATORS

Simulators can offer close-to-reality training in many important safety and operationally related tasks without any physical risk to the ship or trainee. If practical facilities are not available, examiners rely on results from written or oral examinations or a course of training for confirmation of competency. Without any hands-on proof of ability to perform tasks, however, competence can only be assumed but not demonstrated. Even if a student completes a course of training satisfactorily, this does not infer that he/she is competent (by the author's definition – being able to perform a task safely and effectively on demand, any time).

The amended STCW Convention [1] (fully in force 1 February 2002) made specific provision through regulation I/12 for simulator performance standards prescribed in the mandatory STCW Code A for any certificate, to be complied with in the case of:

- any assessment of competency, by means of a simulator
- any demonstration of continued proficiency, by means of a simulator.

If we are going to assess any competence by simulator we need to be able to validate the performance of both the 'assessment system' and the student. Assessment outcomes must be criterion-based and based on real world shipboard operations if confidence in the assessment criteria and ability of the student to achieve the objectives are reliable and valid. In the final analysis, the assessor must consider whether the outcomes are iterative (i.e. repeatable) in nature and thus can be accepted with confidence.

A review of some of the efforts made by researchers and simulator operators to conduct 'assessment' using marine simulators will be useful to the topic.

## 2. A REVIEW OF SIMULATOR BASED ASSESSMENT

Radar/navigation simulators have been in operation for more than forty years, engine simulators for more than thirty years. Early developments with shiphandling simulation started in the late 1960s in Japan and Sweden, followed by shadowgraph nocturnal systems in The Netherlands and UK. Modelboard systems, derived from the airline industry, made their mark in the 1970s in New York and in the Netherlands. Slide projection systems were also in vogue in Germany, Norway, Netherlands and Japan at this time. The emphasis was on training, as many of these systems were constrained to this role by limitations of the simulation technology [2]. The advent of computer generated imagery (CGI) systems, as illustrated by the CAORF system at Kings Point, New York in 1975, proved to be a watershed in the evolution of realistic ship operations on a simulator.

Figure 1: A full mission shiphandling simulator bridge

As simulators became increasingly sophisticated, resulting in a greater functionality and capability, training activities were extended, and the potential of simulation as an assessment tool started to be considered more seriously by researchers and operators. In 1982, Gynther, et al [3] concluded early on that it was difficult to establish reliable objective performance measures in relation to shiphandling. Their view was that precise characteristics of good shiphandling were difficult to define. Reeve and Hurst [4] had some difficulty in quantifying achievement levels in some scoring schemes (on a performed/not performed basis) that they had tested with British Royal Navy officers. Other work by Tin Hlaing [5] further reinforced the view that monitoring performance by scoring was a problem. While agreeing that shiphandling was not an exact science, Muirhead [6] argued that a professional performance carried out in a seamanlike manner could be gauged against a careful selection of criterion-based objectives, supported by the subjective experiences of qualified assessors. His extensive research centred on establishing criteria of performance, against which the seamanlike quality of the manoeuvre, or manoeuvres, could be judged.

Earlier in 1981, Gardinier and Hammell [7] examined the important link between system design and the effectiveness of a training program in achieving training objectives (or performance criteria). They noted that, in any assessment of shiphandling performance, multiple performance measures were called for, which could conflict with each other. This could create a complex monitoring task for an assessor. Hussem et. al. [8] concluded that before performance tests were executed on shiphandling simulators, it was important to establish that the objectives could be fulfilled. STCW Code A-I/12 guidelines on assessment by simulator cover many of these aspects. Another important aspect of simulator assessment is assurance that the skills acquired, under measurable conditions, transfer to the real world. A number of research reports, funded by CAORF [9] , [10], by Reeve and Hurst [4] and Froese[11], supported the view that a positive transfer effect can take place under controlled conditions.

The literature shows little further research activity in the field of simulator assessment between 1984 and 1993, when review of the STCW 1978 Convention commenced. The latter, since its finalisation in 1995, has been a catalyst for further investigations into the use of simulators for assessment purposes. This is not a surprising outcome, noting that the new emphasis of the revised Convention was to require seafarers to demonstrate an ability to perform functions and tasks safely and effectively.

In 1993, Cross [12] argued that the power of the computer could be used to ease the monitoring burden of the instructor/assessor by taking over certain tasks. One role could be that of an objective evaluation of performance, recording actual exercise data and comparing it against predefined performance parameters. This Training and Evaluation Control (TEC) would be much more applicable to a ‘closed’ process, such as main engine operations, than to an ‘open’ navigation process, which is subject to more random variables. In the same year Jenkins [13] concluded that using an engine room simulator as part of an examination system, and relying on the testing and measuring methods of terminal performance (supported by clear objectives), could offer a solution to the problem of assessing skills competency. As seen in Figure 2 below, an assessor is able to follow and record the functioning of the machinery as it responds to the actions of the candidate, while at the same time, being able to observe and record the movements and activities of the candidate (Figure 3). Designs such as this provide the assessor with an excellent environment from which to assess performance.

Figure 2: ER simulator instructor console

Figure 3 ER simulator control room

Meurn and Sandberg [14], who have carried out much pioneering work on simulator training of cadets, developed their first weighted deck watchkeeping marking sheet, containing 47 tasks or criteria, against which performance was to be observed and recorded on a scale of 0-2. With up to four cadets on the bridge, some considerable subjectivity must have remained, but it was a step forward. Schaafstal et al. [15] investigated the feasibility of developing a methodology to conduct proficiency testing in an objective way. Analysing the system, in terms of the functions to be performed, has value for coordinating the activities of system engineering functional descriptions. Task analysis was focused on pilot's and master's tasks with emphasis on perception, information processing, handling characteristics and interface with the workspace. It was noted that the validity of the simulator and scenario would be keys to any successful outcomes, affirming what earlier studies by Hammell [16] had concluded.

Following on from his earlier paper, Cross [17] further described the development of a Simulator Exercise Assessment (SEA) system. In essence, it represented an automatic system that would continuously monitor selected exercise parameters and compare them with criterion values. Any deviations from the norm could be counted and weighted. One advantage envisaged was its potential to provide support to the evaluation process by measuring and recording trends in performance, either of an individual or of a group. The system disciplines the instructor to set clear learning objectives, identify key parameters, set acceptable criterion values for them, and judge their importance. The difference to previous intuitive approaches is that now all is revealed at the debriefing, producing a more objective outcome.

The importance of monitoring and feedback to the assessment process is also illustrated by the increasing interest by other researchers in testing automatic monitoring and feedback systems. Hooper, Witt, & McDermott [18] described a pilot study utilising hypertext and web tools to deliver exercise advice and feedback in electronic format. This has affinity with the expert system of 'Officer of the Watch', for example. Future trials are looking at embedded online assessments. Smith I. [19] described the development of Instructorless Training (ILT), where a trainee can start, undertake and stop the simulator exercises, without any referral to the instructor. Data via printer is then provided to allow the student to compare personal performance against a model response to the exercise. There is a need to recognise, however, that such approaches have limitations as to how the final judgement about the reliability and validity of the performance against the set criteria is to be made. While technology can be helpful in giving a trainee immediate feedback, the assessor cannot become disengaged from the assessment process entirely. To what extent do cognitive, affective and psychomotor skills form part of the measure of performance? Visual perception by an experienced assessor must be mandatory, in most cases, when determining competency to perform tasks or functions. There is a danger that technology becomes the yardstick rather than being seen as another supportive tool to the evaluation process.

Continuing their work with simulators, Meurn and Sandberg [20] further refined their approach to the assessment of watchkeeping skills. In association with other US academies, a list of Knowledge, Understanding and Proficiencies (KUPs), drawn from Table A-II/1 of the STCW Code, was created in 1999. Several KUPs have been designed to be assessed during the Kings Point Bridge Watch Standing Course. The mariner assessment consists of five steps, namely: identifying test objectives, determining test modes, specifying test conditions, developing performance measures and standards, and preparing test materials. A performance measure indicates how a trainee's performance will be observed and recorded. A performance standard represents the level that is established as acceptable. Drawing on their experiences, the authors observe that the effort of seeking more objectivity results in some loss of validity, their view being that the use of experienced mariners, trained in simulator assessment techniques, is the best way of minimizing such loss of validity.

In the aftermath of the amended STCW Convention in 1997, the United States Coast Guard (USCG) recognised the need to provide guidance to assessors on assessment techniques, if the latter were to affirm the ability of mariners to demonstrate practical skills and knowledge. In 2000, the USCG National Maritime Centre (through the USCG Research and Development Centre) sponsored a research project that was designed to build on earlier efforts that had produced a specification for a Performance Based Assessment (PBA) method for creating mariner assessments [21]. The intention was to further refine the PBA method (which also addresses the conduct of assessments), and provide practical support to assessors in the USA. Through cooperative workshop activity, a manual for assessment developers was created, and several sample assessment procedures were developed.

Kobyashi [22], who has also been very active in developing assessment methodologies, emphasised the importance of the relationship between required behaviours, simulator functionality and a clear analysis of objectives. Arms [23] has described how the full-mission simulator was used to assess the shiphandling competency of trainee pilots, as part of the process of selecting applicants for the Bay of San Francisco pilot trainee program. Four trained assessors observed the outcomes, using a scale of high, acceptable, ineffective or fail to judge candidate performance. The role played by the trained assessors was considered critical to the validity and success of the program. It should be noted that the time to prepare and administer the program was spread over several years.

From a quality assurance point of view, an interesting approach is the requirement by the USCG authority that an expert assessor will be qualified as a designated examiner. Muirhead [24] had earlier emphasised that maritime administrations should take more positive steps to train and use experienced personnel, both at sea and ashore, as assessors in the wake of STCW. Again, the USCG project emphasises the need for specification of performance measures, performance standards and proficiency criteria, if the outcomes are to be reliable and valid. A particularly vexing question that arose centred on the level of detail required to assure reliability and validity. Another difficulty is obtaining consensus from a group of experts conducting validation of processes and procedures, where personal experiences and subjectivity may hold sway at the expense of objective criterion approaches. Similar problems were experienced in earlier research by Muirhead [6] when shiphandling performance of trainees was validated using a criterion-referenced approach to establishing competence.

A USCG Research and Development (R&D) centre report [25] released in January 2001, described a process that would evaluate the "capability of simulators to support performance based assessment of mariner proficiencies"**.** It emphasised again that if a simulator lacks the design features to cover the functionalities being assessed, then any performances measured must be both unreliable and invalid. The Det Norske Veritas simulation certification standards [26] are also designed to minimise the risk of such situations arising. Among a number of recommendations made in a final USCG R&D report of May 2001 on "conducting mariner assessments" [27], perhaps the most important one was the following statement:

> Those in the industry who are responsible for training and assessment of mariner proficiency in academies, training schools, and shipping companies should make use of the method and the materials presented here as a guide for their own development of assessment procedures
>
> USCG: CG-D-02-01, p.vi

Should it be so strange an idea, to require those charged with the task of assessing the performance of mariners, to be themselves licensed as being fit and proper persons to conduct such assessments?

## 3. SIMULATOR CLASSIFICATION AND FUNCTIONALITY

The International Marine Simulation Forum (IMSF) has attempted, for a number of years, to classify simulators by type and relate that classification to the functionality needs for assessing competence. In 1993, a working group of IMSF submitted to IMO, proposals for a simulator classification system based on three types of simulator on four levels of functionality. This was further refined into four classes of simulator by the consultants to the STW sub-committee. However, for various technical reasons, the proposals were not incorporated into the revised STCW Convention. Recognition that agreement on this and other technical issues would not be reached easily led to other approaches being made.

Cross [28] proposed creating a model linking the functional requirements of the Convention to the classification work done by IMSF. In a following paper, Cross (2000) [29], noting that STCW identified the learning objectives required to be met for the different levels of officers, functions and subjects, considered how this could be reconciled with the requirements of regulation I/12 for demonstrating competence through approved simulator training. Any competence based assessment system has measurable standards of performance and under quality assurance requirements of STCW these must be approved by the administration. Where a simulator is being used

to assess any competence then it too must meet the performance standards laid down.

Cross rightly pointed out that not all manufacturers' simulators of a described type had the same facilities or capabilities, and thus could not be relied upon to measure the same functions or tasks. Taking this functional approach as a basis, Cross asked the question; "which functions can be performed and demonstrated on what level of simulator"? Taking knowledge of the collision regulations (Colregs) as an example, Cross undertook the considerable task of examining the relevant Code A tables to identify which functions and tasks could be completed using a simulator. This information was then correlated against the relevant collision rules and IMO model courses. For example, while some maritime authorities (UK, Norway) have recognised certain desktop simulation products as being suitable to cover certain training functions at watchkeeper level, they do not recognise them as systems capable of offering functionality for shiphandling skills, even though the systems provide manoeuvrable ship models.

Table 1 Bridge simulator applicability

| Type of Simulator | Colreg Rules | STCW 95 | IMO Model Course |
|---|---|---|---|
| Full Mission | 5-10<br>12-19<br>20-31<br>34-37 | Table A-II/1 item 1-6<br>Table A-II/2 item 1-8<br>Table A-II/3 item 1-5<br>Section A-VIII/2<br>Section B-V/3 | 1.07<br>1.09<br>1.08<br>1.22 |
| Multi-task | 5-8, 10 | Table A-II/1 item 1-6<br>Table A-II/2 item 1-6<br>Table A-II/3 item 1-5<br>Section A-VIII/2 | 1.07<br>1.08<br>1.09<br>1.22 |
| Limited task | 6-10<br>19 | Table A-II/1 item 3-6<br>Table A-II/2 item 6<br>Table A-II/3 item 3-4 | 1.07<br>1.08<br>1.09 |
| Single task | 19 | Table A-II/1 item 3 | 1.07 |
| CBT | 23-31 | - | - |

Source: Cross [29] p. 20

Taking up the challenge, Det Norske Veritas published a set of standards for the certification of maritime simulator systems, in 2000. Using a configuration of four classes of simulator and four function areas, the standards focus on providing users and administrators with the framework to establish that defined training and/or assessment objectives can be fulfilled. While the standards target the simulator system, being just one of three key aspects (the others being the instructor/assessor and the approved program), they provide an excellent link between simulator capability, assessment objectives, functions and tasks and STCW competencies. As such, the approach to assessment by simulator has been enhanced and clarified [30].

To summarise, a simulator must have a functionality capability to meet the training or assessment objectives, otherwise performance will be neither reliable nor valid. In the latter case, any assumptions made regarding the transfer of skills to the real workplace, will also be invalid.

## 4. STCW AND ASSESSMENT

### 4.1 Assessment procedures

STCW Code A (A-I/12) [1] provides assessors with a comprehensive list of assessment procedures that are to be followed. Compared with training, the emphasis is somewhat different. Candidates must know what the performance criteria is and be clearly briefed on the tasks and/or skills by which competency is to be determined. Assessors must ensure that assessment criteria have been selected with optimum objective measurement and evaluation in mind, keeping subjective judgements to a minimum. Assessors are also warned to use scoring or grading methods with caution, until such methods have been validated.

As seen earlier, a number of studies have looked at the use of scoring methods to establish competency. McCallum [31] and Smith C. [32] carried out further valuable work in this area. Using weighted performance scoring, Muirhead [33] conducted a series of simulator instructor training courses in the Netherlands between 1999-2004, where performance criteria were judged on the basis of the competence factors being either essential (E) (score range 0-5) or desirable (D) (score range 0-3). Scores were awarded against a set of qualitative descriptors. A number of specific quantifiable parameters or limits were set against the assessment criteria. The purpose was to ensure that the outcome was measured against both quantitative and qualitative criteria, to reduce reliance on more subjective judgements made on the basis of 'good seamanship'. The latter observations, however, still continue to form a very important element of the overall judgement.

Performance outcome was then measured using the following weighted formula:

Competency Score = Number of E factors x 5 + Number of D factors x 3 x 70/100 (i.e. set at 70%)

Example: there were 11 E factors and 5 D factors, the Competence Weighted Score = (11 x 5) + (5 x 3) x 70/100= 49

A candidate with an Actual Performance Score above 49 would be considered competent, provided no Essential items had been failed (i.e. obtained a score of less than 3 for an E item).

The outcomes from these courses indicate that the mix of quantitative measures, qualitative observations, reports and the use of a peer review process at debriefing has helped to reduce the subjectivity of the judgement process, and provided instructors with increased confidence in their ability to use this extended range of tools to better assess competence in a more reliable and valid way. This is even more marked when using the engine room simulator. However, such procedures are more time consuming and labour intensive for the assessor(s).

### 4.2 The assessor

The STCW Code A is quite specific in relation to simulator assessment and states, inter alia:

*Assessment of competence*

> Any person conducting in-service assessment of competence of a seafarer ... shall:
>
> .5 if conducting assessment involving the use of simulators, have gained practical assessment experience on the particular type of simulator under the supervision and to the satisfaction of an experienced assessor.

STCW Code A: section A-I/6, part 6

Global surveys of MET institutions by Muirhead [33] have indicated that there is a hesitancy to use simulators for assessment purposes. EU research projects MASSTER, METNET, METHAR & EASTMET also showed a marked reluctance amongst instructors to use simulators to assess the performance of individuals. There is strong circumstantial evidence that lack of knowledge of technique is one factor inhibiting instructors from doing more.

Although section A-I/6 of Code A of STCW clearly requires an assessor to have had practical assessment experience on a similar simulator before conducting assessment, there is little evidence outside of the USA, Netherlands and the UK of Parties ensuring that both present and new generations of simulator instructors have been trained to conduct the assessment function. It is not surprising, therefore, that there is such a low level of activity globally.

Evidence from the most recent research activity is that clear guidance on the procedures and processes to be used to assess the competency of seafarers using modern technology is now available. A convincing case can be made to require all simulator instructors to undertake formal mandatory training in assessment techniques.

## 5. SOME KEY QUESTIONS AND ANSWERS

### 5.1 Use of Simulators: do they provide an experiential training medium only?

Can the competence of individuals to perform functions and tasks safely and effectively by simulator ever be more than achieving satisfactory training outcomes, i.e. acquiring short-term experience of situations? Many would argue that, while training can effectively prepare individuals for workplace experience, it does not provide them with the degree and range of experience to deal with workplace reality. Experience can only come with further exposure to real life. The counter argument is that no individual will ever experience all possible workplace events, and simulation can help overcome many of these omissions in a safe training environment. Most simulator courses are providing value-added training at various levels, but not determining directly the competence of individuals to perform in the workplace.

### 5.2 Why so little formal assessment of competence by simulator?

Despite the previously described assessment activities, most simulator programs are given recognition only for their ***training*** outcomes as part of the overall approval process by the Maritime Administration of MET course programs. Most instructors, in practice, do not have the time available to conduct formal assessment of individual performance at watchkeeper level. Instructors are unlikely to be approved assessors or even have access to one. Even less likely is assessment of competency or ongoing proficiency of mariners at Chief Mate/Master level.

It is a disturbing fact that with the majority of MET programs delivered through a front-ended format today, many administrations award certificates of competency at these higher levels solely on the evidence of further time served at sea. They have no practical evidence that mariners are competent to perform specified functions at these higher levels.

For example, one requirement of STCW Code A, Table A/II (Masters and Chief Mates on ships of 500 GT or more) is to manoeuvre and handle a ship in all conditions. Why is this loophole in STCW not being closed by at least a requirement for further mandatory simulator training and assessment? Generally the Parties in particular (i.e. administrations), as well as shipping companies and institutions, are reluctant to commit resources to a safety issue that lays quietly dormant!

**5.3 Are simulators complying with STCW standards?**

IMSF and EU surveys of simulator training centres show that, in the great majority of cases, STCW Code A performance standards are being applied correctly and systems are being used for STCW functionalities within the technical capabilities of the simulator systems.

**5.4 Is there a lack of approved assessors?**

Formal courses of training in assessment techniques for simulator assessors are the exception rather than the norm, formal courses are known to be available in the USA, Netherlands and UK. There is a lack of a global standard for such training. It is therefore not surprising that formal assessment of competency or ongoing proficiency allowed under STCW regulation I/12 is so limited or non-existent. Few administrations have established a process for approving persons as simulator assessors.

**5.5 Does simulator based competence assessment lack reliability and validity?**

As illustrated earlier, there has been considerable research undertaken into the assessment aspect. It is recognised that there is much diversity of approach, but all are agreed on several points. Reliable objective performance measures and standards must be specified, simulator design features and systems must cover the functionalities being assessed, and the assessor must be trained in assessment techniques and have experience of conducting such assessments on a simulator. At present assessment methods lack clarity and uniformity.

**5.6 Can IMSF do anything to achieve a more global standard?**

IMSF can play a leading role here by establishing cohesive assessment standards for various functionalities and tasks for its members. Without this, administrations are unlikely to accept simulator assessment as a valid and reliable measure of competence as allowed for under STCW regulation I/12.

**5.7 Is it realistic to expect marine simulators to be used for assessing mariners' ongoing proficiency in the same way as the airline industry?**

While the airline industry's simulator-based type rating assessment is made easier by the limited number of aircraft types on the market, there is no reason why a system of assessment of senior officers for on-going proficiency by marine simulator could not be put in place, particularly if IMSF established some form of globally recognised standard or procedure. There would be much to be gained from a safety point of view in the handling of many of the specialised vessels operating today. The drawbacks are that good preparation takes time and using simulators for individual assessment is a labour intensive and relatively expensive process.

## 6. CONCLUSION

Reviewing the use of technology to assess the competence of seafarers in the last twenty years shows that developments have ranged from the creation of quantifiable criteria of performance, weighted scoring methods, an automated comparative approach between weighted exercise parameters and set criterion values, the use of an iterative process of review, discussion and consensus by subject matter experts, the development of a model for Performance Based Assessment, and the use of weighted criterion reference scores.

The efforts of IMSF in the 1990s, and of Det Norske Veritas in producing in 2000 a classification of simulation systems based on the link between simulation functionality and capability, and competency requirements of the STCW Convention, provide another very important step in creating a reliable and valid assessment process with technology.

It can be stated, with some confidence, that positive measurement of skill acquisition can be made on simulators, subject to the provisos that the criteria for performance measurement is clearly specified, the simulator provides an appropriate operating environment, the assessor is properly trained and experienced, effective recording and monitoring equipment is available, and that the assessment functions and tasks selected relate to the real world of ship operations.

The documentation from the research carried out by the United States Coast Guard on the training of assessors is fully available on the Web. It is recommended that this excellent database be tapped by IMSF members as a basis for creating a more common approach to assessment by simulator, supporting it with the experiences gained in the Netherlands and UK and other known sources.

There is a clear omission in STCW implementation processes that fails to ensure that senior officers equally have demonstrated their competency or ongoing proficiency at higher certificate levels. The potential use of simulators to assist in this regard remains largely untapped. There is, in the author's view, a clear compelling case for IMO to consider changing the status of performance standards for

other simulators laid down in Code B I/12 from guidance only to a more mandatory requirement in Code A.

A more proactive approach is needed by the IMO, industry, administrations, institutions and simulator operators, to ensure that standard assessment measures are put in placed and implemented and that sufficient assessors are properly trained for the task. Only then is it reasonable to expect that performance measurements of competency and ongoing proficiency by the use of marine simulators will be seen as reliable and valid, and thus an acceptable method of increasing confidence in standards of competency at all certificate levels.

## REFERENCES


[1] International Maritime Organization (1995). *Final Act of the Conference of Parties to the International Convention on Standards of Training, Certification and Watchkeeping for Seafarers*, 1978. London: IMO.

[2] Muirhead P.M. (1985), A report on the specification, installation and training profile of a new full mission shiphandling simulator, IMSF Meeting, 1-2 March 1985, Kings Point, New York.

[3] Gynther, J.W., Hammell, T.J., Grasso, J.A. & Pittsley, V.M. (1982), *Simulators for mariner training and licensing: Guidelines for deck officer training systems*. Report No 50-8004-03, CAORF, New York.

[4] Reeve, P.E. & Hurst, W.E. (1984), Evaluation of simulator-based training in visual navigation for junior Royal Navy officers and its transfer to sea, *Proceedings of 3rd International Conference on Marine Simulation, Marsim 84*, IMSF, Rotterdam.

[5] Tin Hliang, U. (1984), Assessment of officer watchkeeping performance using a ship simulator, University of Wales, Cardiff.

[6] Muirhead, P.M. (1987). The assessment of shiphandling skills on the ship simulator. *Proceedings of 4th International Conference on Marine Simulation (Marsim 87).* 22-24 June, 1987. Trondheim, Norway.

[7] Gardinier, J.S. & Hammell, T.J. (1981), Shiphandling simulator standards and training technology, *Proceedings 2nd International Conference on Marine Simulation, Marsim 81*, IMSF, New York.

[8] Hussem, J., de Froe, C., Dijkhuizen, H.A., Luchters, J.H.M. (1984), Training and testing of pilots and VTS operators on simulators in the Netherlands. *Proceedings of the 3rd International Conference on Marine Simulation, Marsim 84*, IMSF, Rotterdam.

[9] Miller, W.C., Saxe, C., D'Amico, A.D. (1982), *A preliminary evaluation of transfer of simulator training to the real world*, CAORF, New York

[10] Multer, J., D'Amico, A.D., Williams, K., Saxe, C. (1983), *Efficiency of simulation in the acquisition of shiphandling knowledge as a function of previous experience*. Report No 52-8101-02, CAORF, New York

[11] Froese, J. (1987), Simulation v real ship experience: An investigation into transfer training effects, Report, AMC, Launceston

[12] Cross, S.J. (1993). Objective assessment of Marine Simulation training, *21st Century Technology Conference*, Southampton Institute, Southampton

[13] Jenkins, M.E. & Theedom, R.C. (1993), Use of simulators as a training and examination tool, *Proceedings of the 6th International Conference on Marine Simulation (Marsim 93).* IMSF, St.John's, Newfoundland

[14] Meurn, R. & Sandberg G.R. (1993). Ship simulator bridge and control station requirements from an instructor's point of view. *Proceedings 6th International Conference on Marine Simulation, Marsim 93*, IMSF, St John's Newfoundland.

[15] Schraafstal, A.M., van Maanen, W. Ph., Wolters, A.J.W., Lems, J.C. (1993), Training, Proficiency Testing, and Licensing of Masters of Sea-Going Vessels Using a Full-Mission Simulator, *Proceedings of the 6th International Conference on Marine Simulation (Marsim 93).* IMSF, St.John's, Newfoundland

[16] Hammell, T. J. (1981). The training device is more than a simulator. *Proceedings of 2nd International Conference on Marine Simulation (Marsim 81).* New York.

[17] Cross, S.J. (1996). Methodology for bridge simulator skills assessment. 7th International Conference on Marine Simulation (Marsim 96). 9-13 September, Copenhagen, Denmark.

[18] Hooper, J.B. Witt, N.A.J. & McDermott, A.P. (2000), Automatic Student Feedback and Navigation Simulation, *11th International Navigation Simulator Lecturers' Conference (INSLC)*, Kalmar, Sweden.

[19] Smith, I. (2000), Instructorless Training, *11th International Navigation Simulator Lecturers' Conference (INSLC)*, INSLC, Kalmar, Sweden.

[20] Meurn, R. & Sandberg, G. (2000). The assessment of Watch Standing Skills using a Bridge Simulator, *Proceedings 11th International Navigation Simulator Lecturer Conference*.(INSLC) 14-18 August 2000, Kalmar, Sweden.

[21] McCallum, M.C. & Smith, M.W. (2000) Evaluating simulators for use in assessment of mariner proficiencies. *Proceedings 8th International Conference on Marine Simulation*,

*Marsim 2000*, IMSF, May 8-12, Orlando, Florida.

[22] Kobyashi, H. (2003). *MARSIM 2003 International Conference on Marine Simulation & Ship Manoeuverability, Vol. 1. August 25-28, Kanazawa, Japan.*

[23] Arms, P.B. (2004). Competency assessment of pilot trainee applicants using a full mission bridge simulator. *13th IMLA International Conference.* St Petersburg. 14-17 September 2004.

[24] Muirhead P M (1996) The amended STCW 78 Convention and the use of marine simulators: Impact on the training and assessment of seafarers by simulators and the qualifications of simulator instructors and assessors, *9th International Navigation Simulator Lecturers' Conference*, August 19-23 1996, Vancouver, Canada.

[25] United States Coast Guard: *Evaluating simulators for assessment of mariner proficiency*. CG-D-05-01 Final report January 2001. USCG Research & Development Centre, Groton, CT.

[26] Det Norske Veritas, (2000). Standard for certification of maritime simulator systems, No 2.14. Oslo: DNV.

[27] United States Coast Guard: *Conducting mariner assessments*. CG-D-02-01 Final report May 2001. USCG Research & Development Centre, Groton, CT.

[28] Cross, S.J. (1998). The Functional Approach to Simulator Classification, *IMSF Annual General Meeting*, IMSF, Sydney.

[29] Cross, S.J. (2000). Simulator specifications in relation to the STCW competences. *11th International Navigation Simulator Lecturers' Conference (INSLC)*, INSLC, Kalmar, Sweden.

[30] Cross, S.J. & Olofsson, M. (2000), Classification of maritime simulators, the final attempt: introducing DNV's new standard, *8th International Conference on Marine Simulation (Marsim 2000)*, IMSF, Orlando, Florida.

[31] McCallum, M. C. (1996). Criterion-based assessment of marine simulator operator performance. *Proceedings of 1st International Conference on Marine Simulation towards Safer Seas and Cleaner Oceans*. Alexandria, Egypt.

[32] Smith, C. (2000), Shiphandling on Simulated Seas, *11th International Navigation Simulator Lecturers' Conference (INSLC)*, INSLC, Kalmar, Sweden.

[33] Muirhead, P.M. (2003), A study of the impact of new technology and teaching methodologies on global maritime education and training into the 21st Century, Published Ph.D. Thesis, Curtin University of Technology, Perth, Australia.

## AUTHOR'S BIOGRAPHY

**Peter Muirhead** was the Inmarsat Professor of Maritime Education and Training at the World Maritime University (WMU) for eleven years before retiring in 2004. Prior to that he was surveyor and examiner of Masters and Mates in Sydney for over 5 years and a foundation senior academic at the Australian Maritime College between 1979-1993. In addition to being a qualified Extra Master Mariner, he holds both Ph.D and M.Sc degrees in the use of new technology and methodology in maritime education and training and has published widely on these issues. He has been involved in the use of marine simulators for over 30 years and remains an honorary member of the IMSF. As a consultant, he assisted the IMO between 1993 and 1996 in the review of the STCW 1978 Convention and in the subsequent global implementation programs, and led the development team that produced the first IMO disc-based e-learning program in 2004. He continues to support the WMU as Adjunct Professor.