Haavardtun, P., Mallam, S., Møller, L. & Høifødt, F. A. (2025). Exploring students' view on summative assessment: Case Study in Marine Engineering Education. I P. Haavardtun & L. I. Magnussen (Red.), *Læring i maritim næring* (s. 131–147). Fagbokforlaget. DOI: https://doi.org/10.55669/oa500107

Kapittel 7

Exploring students' view on summative assessment

Case Study in Marine Engineering Education

Per Haavardtun, Stephen Mallam, Lasse Møller and Fred Arne Høifødt

Abstract: The education of marine engineers is a practical-oriented program typically utilizing differing training simulations and workshops to replicate real work practice. The summative assessments implemented tend to be conducted with traditional assessment methods, such as written closed-book or oral testing. This paper describes the students' view on three different practical assessments using a marine engine system training simulator (Big View). The purpose was to better understand students' perceptions, experiences, and preferences between written and practical assessment frameworks in a course with a very practical-oriented learning objective where system understanding, and cause and effect implications are essential.

This research collected data from 14 Bachelor of Marine Engineering students using written questionnaires focusing on the students' feelings and opinions. The results found that students preferred the assessment method most resembling their perception of the learning objectives in the course. In this case, the objective focused on system understanding and skills. A practical assessment, in comparison to traditional knowledge-based written exams, was preferred as students reported that they were more accurately measured on learning objectives and skills. This assessment approach was found to be increasingly favoured as the complexity of the simulator scenario increased.

The students felt the test had been unbiased even though they saw the possibilities for bias. The students' conclusion was based on the teacher's ability to be rigid in the tests and the transparency on how they would assess the work. This shows that focus on fairness in the execution of a practical examination assessment is more important than the need for a total objective assessment regime.

Keywords: student engagement, problem-based assessment, summative assessment, non-technical skills

Introduction

Maritime training simulator technology has become increasingly sophisticated in recent years. Many simulators now offer highly realistic and immersive experiences for trainees. This has helped to make practical simulator training and trainee assessment an increasingly important part of maritime training programs (Sellberg, 2017). Simulator-based training has been found to improve the knowledge and skills of mariners (International Maritime Organization, 2010). It also enhances problem-solving skills and the trainees' ability to perform under stress (Oser et al., 1999).

The ongoing imperative to develop novel methods that enhance student learning is crucial. It ensures the integration of the preferences of the new generation of students and their inclination towards digital solutions within the educational framework of universities. This integration is pivotal for presenting the requisite skills, knowledge, and understanding that students must acquire to effectively function within their intended professional spheres.

Typically, all university courses include some form of summative assessment to assign individual grades based on student performance. Courses adhering to stringent regulations, such as maritime studies governed by the International Maritime Organization (IMO), often employ stricter assessment methods. These usually involve written closed-book exams. However, such exams make it challenging to measure practical competence. This is often assessed through formative evaluations using real machinery or simulators during the course.

Research has shown that students' perceptions of how the summative assessment will be conducted can have a significant impact on their approaches to learning and studying (Struyven et al., 2005). Knight (2004) found that students reported preferring individual assessment. Van de Watering et al. (2008) found that students who believe that assessment was fair and aligned with the goals of the course were more likely to engage in deep learning strategies, such as elaboration and self-explanation (Van de Watering et al., 2008). When students perceive that assessments are aligned with the goals of the course, it challenges them to take responsibility for their own learning. Active learning processes connected to the goals of the course make it more likely that students will adopt a mastery approach to learning, characterized

by a focus on understanding and long-term retention of material (Pérez-Sabater et al., 2011).

Students with a growth mindset (i.e., the belief that their intelligence can be developed) were found to be more likely to respond positively to an assessment (Cavanagh et al., 2018). Growth mindset students approach assessment as an opportunity for learning and development. In contrast, students with a fixed mindset (i.e., the belief that their intelligence is set in stone) are more likely to view assessment as a threat (Cavanagh et al., 2018). Thus, it can be argued that students who view summative assessment as a natural milestone in their competence building will be positive about the assessment.

Marine engineering students have a very practically oriented education program. There is a clear link between the educational curriculum, courses, and assessments and the real-world skills and competencies required by marine engineers once they leave the university setting. Thus, this curriculum structure and its goals place students more towards a growth mindset, where they challenge the system to give them the possibility to grow.

Purpose & Scope

Currently, our university is engaged in the development of a comprehensive initiative. Maritime simulators will be, together with miscellaneous software solutions, both commercially available and developed internally, integrated into an extended simulation. This initiative involves the integration of serious simulation games from the different maritime professions taught at our university within an extended simulation framework spanning several semesters. The overarching objective is to provide maritime students, enrolled across diverse maritime programs, with an in-depth comprehension of the intricacies and functioning of the maritime industry. Furthermore, the initiative seeks to elucidate how the various professions within the maritime domain, including nautical, marine engineering, shipping, and logistics, coalesce within this industry.

This paper looks at possibilities to include summative assessment from given courses as part of the extended simulation. This means that a written exam, as the traditional method of assessment, will be replaced with a practical assessment using available simulators. Assessment of practical competence with training simulators is an important aspect of evaluating the effectiveness

of training programs, including the maritime domain. However, it is essential to have a valid and reliable assessment system in place to ensure accurate and repeatable measurement. The purpose of this paper is to investigate a new and more practical-oriented assessment protocol within a Bachelor of Marine Engineering program. A new assessment protocol was designed and implemented with three marine engineering simulation activities. This study looks to better understand student and instructor perceptions of simulator training assessment. This investigation and data will provide insights and feedback into a new assessment approach within the marine engineering program and contribute to further development of practical summative assessment in practical-oriented subjects, and refinement of more objective and transparent assessment techniques.

Description of the case study course and its teaching methods and approach.

The course "Marine Machinery and Motor Propulsion," offered within a Bachelor of Marine Engineering program at a Norwegian university, serves as our case study. This course is practically oriented but necessitates a strong foundation in mathematics and physics for students to comprehend the systems and processes involved.

Several of the course's defined learning outcomes emphasize the acquisition and demonstration of skills related to safe operation and correct procedures in an engine room. For instance, students are expected to "maintain a safe engineering watch" and "lead and build the engine room team, as well as conduct safe work and operational practices." These skills are developed through a combination of theoretical instruction and practical sessions in an engine room simulator.

As a mandatory course with compulsory attendance and required passing grades for eligibility to apply for a marine engineering certificate, students' knowledge must be evaluated through summative assessment. This ensures legitimacy (Kulasegaram & Rangachari, 2018) and qualifies them for standardized international certification as marine engineering officers (International Maritime Organization, 2010). Additionally, course improvement, teaching effectiveness, and the overall program's evaluation are influenced by the results of this summative assessment (Kibble, 2017). Therefore, it is crucial to align

the assessment method with the learning objectives agreed upon by both students and instructors.

Traditionally, this course has utilized a written closed-book exam as the summative assessment. However, this method may not be optimal given the course's practical learning outcomes. The non-technical skills (NTS) aspect of the course, including interactions between personnel in the engine room and engine control room, cannot be adequately assessed through a closed-book written exam. It does not allow students to demonstrate or instructors to evaluate these skills effectively.

This case study is based on the premise that in a practically oriented program like marine engineering, learning involves a combination of procedural and NTS knowledge. Instructors collaborate to assess students' learning outcomes and provide them with an understanding of their strengths and weaknesses through practical demonstrations and testing. The challenges of subjectivity, biased assessors, and performance standard variations make this difficult. Thus, this study aims to develop and test a practical assessment within the marine engine room simulator, focusing on objectivity and unbiased evaluation.

The questions posed are:

- Do students have any preferences for assessment methods based on the learning outcomes of the course?
- How do students assess equality, fairness, and objectivity in the grading of a practical exam?

Research Methodology

The study utilised a mandatory formal assessment for all students, graded on a pass/no-pass basis. This assessment aimed to verify that students possessed the necessary practical competence and knowledge to qualify for the final exam. Students conducted the assessment in pairs. One student operated the engine room, and the other managed the engine control room (see Figure 7.1). A total of 14 students participated, forming 7 pair groups. The data was

gathered in the fall of 2022. Each session was observed by two teachers who facilitated the scenarios and continuously assessed the students. Observations were documented, and grading was completed at the end of each session.

Figure 7.1

Marine engineering training simulator: one student (right, foreground) manning the engine control room, one student (left, background) manning the engine room.

Each session comprised three distinct scenarios, each with an assessment sheet designed to evaluate specific outcomes.

Scenario 1 focused on assessing students' foundational knowledge through a checklist of 33 procedural steps required for the safe initiation of the main engine. To pass, students needed to complete 90% of the checklist without missing any critical elements.

Scenario 2 evaluated students' ability to integrate various knowledge domains and understand the interconnectedness and interactions within the system. The assessment sheet for this scenario listed 17 action points that had to be completed before transitioning the engine room to an unmanned state. Failure criteria included triggering an alarm within the first 30 minutes of unmanned operation or omitting more than 40% of the action points.

Scenario 3 assessed students' comprehensive understanding of the entire engine system and its interactions with auxiliary systems. This scenario introduced simulated errors, requiring students to identify and describe the consequences and propose corrective measures.

The primary objective of these simulation exercises was to ensure fair treatment of all students during testing and assessment, thereby reducing biases. The skills targeted for learning in this course included both procedural knowledge and non-technical skills (NTS). Instructors collaborated to assess students' learning outcomes, providing real-time and post-hoc feedback on their performance during the simulator scenarios. However, significant deficiencies were noted, such as the lack of explicit frameworks, subjectivity, biases, and unreliability in the assessment of NTS in simulator training.

One week after the user case, students were asked to complete a paper-based questionnaire during class. All students who participated in the user case filled out the questionnaire. The survey employed a 5-point Likert scale, where respondents rated their level of agreement from 1 (complete agreement) to 5 (absolute disagreement). See amendment 1 for an example of the questionnaire. The survey included five main questions aimed at capturing students' views on the assessment paradigm, as shown in Table 7.1.

Following these general questions, the survey included four specific questions for each of the distinct marine engineering simulation scenarios, detailed in Table 7.2. These questions were consistent across all three scenarios to allow for comparative analysis. Additionally, students were given the opportunity to provide written comments explaining their choices, thereby enriching the qualitative insights obtained from the survey responses.

Findings

Students' Views on Assessment

Table 7.1

Survey questions asked and student responses

Question	Average	Standard deviation
I like practical [summative] assessment better than a written one	1.62	0.49
I show my knowledge better in a practical assessment than a written one	1.62	0.49
Practical assessment is more demanding than a written one	3.00	1.04
I mean the result on a practical assessment is more correct than a written one	1.85	0.95

The students were surveyed about their preference for practical assessments over written evaluations. The mean score for this question was 1.62, indicating a preference for practical assessments. Additionally, when asked if they believed they would perform better in a practical test compared to a written exam, the average score was again 1.62, with 1 representing total agreement. These responses suggest that the students have a practical orientation, showing a genuine interest in acquiring practical skills essential for their jobs, beyond just theoretical understanding.

When evaluating the perceived difficulty of practical versus theoretical assessments, students gave an average score of 3.0 (SD = 1.04). This question yielded inconclusive results, and further empirical investigation is needed to understand the underlying reasons. It is noteworthy that the concept of practical assessments in a group setting is relatively new to the students, contrasting with their familiarity with traditional evaluation methods.

Students were also asked about their perception of the accuracy of results from practical tests versus theoretical ones. The average score of 1.85 (SD = 0.95) supports the hypothesis that students believe practical assessments are more effective in measuring proficiency in practical skills. This

suggests that, according to the respondents, practical assessments provide a more authentic and accurate measure of their practical competencies compared to theoretical evaluations.

Further questions were directed towards the students' perceptions of the fairness of opportunities, their confidence in performing well, their belief in equal treatment by instructors, and their understanding of scenarios and associated assessment sheets in terms of procedural transparency. Additionally, the possibility of subsequent evaluations if necessary was considered. Consistency was observed across all three scenarios, allowing students to provide additional comments alongside their selected responses.

Table 7.2

Students' view on the scenarios used in the assessment

Question	Average	Standard deviation
1. "This checklist is a good way to show my knowledge."		
Scenario 1	2.00	1.04
Scenario 2	1.57	0.84
Scenario 3	1.43	0.84
2. "I believe that this method ensures equal opportunities for all students."		
Scenario 1	1.93	1.07
Scenario 2	1.71	0.91
Scenario 3	1.71	1.12
3. "I believe that this method reduces the possibility of treating students differently."		
Scenario 1	2.07	1.00
Scenario 2	2.21	0.97
Scenario 3	1.93	1.18

Question	Average	Standard deviation
4. "I believe this method is transparent with the possibility of verifying the assessments."		
Scenario 1	1.77	0.80
Scenario 2	1.77	0.83
Scenario 3	1.92	1.04

Inquiring about the efficacy of a checklist-based scenario as a demonstration of knowledge in Scenario 1, the mean score was 2.0. Two students expressed disagreement. Critiques from dissenting students highlighted concerns such as a perceived emphasis on memorisation over comprehension. They also asserted that, in practical situations, reliance on checklists is contingent upon uncertainty.

Scenario 2 yielded an average score of 1.57, with one student dissenting. Notable comments pointed out the challenges of memorising intricate system interactions and the potential for triggering alarms despite adherence to the checklist. The necessity for a comprehensive understanding of fundamentals emerged. This was exemplified by a dissenting student's assertion that while knowledge was demonstrated, the sheer volume of information made it susceptible to lapses.

The third scenario garnered an average score of 1.43, with one student expressing dissent. Comments underscored the imperative of possessing substantive knowledge for effective response. They emphasised the necessity for broader and deeper learning compared to the preceding scenarios. Overall, the students conveyed a positive inclination towards this assessment format. They asserted that it facilitated the exhibition of knowledge and enabled instructors to conduct accurate comparative evaluations.

Notably, a discernible pattern emerged wherein students increasingly favoured the novel assessment framework as the complexity of simulator scenarios and corresponding evaluations escalated. This trend suggests that assessment questions requiring the integration of knowledge for problem resolution are well-suited for practical tests conducted within the context of authentic, real-work simulators.

Discussion

This study investigated a practical-oriented course where holistic understanding and the ability to connect different basic knowledge is the main goal. Traditionally, the final exam has been a written closed book assessment, with sensors considering the results and giving a grade. This is primarily due to the need to test students in mathematical calculations, which is also an important basic knowledge needed to understand the whole picture. However, it emphasises shallow learning, not aligned with a course where holistic understanding and problem-solving are of vital importance.

Students' view on equal opportunities and reduction of biases

Ensuring equal opportunities for students to demonstrate their abilities and reducing biases is crucial. Derived from students' general comments on potential enhancements, several key insights have emerged, guiding the ongoing refinement of the assessment framework. One notable observation pertains to the varying proficiency levels among students, as highlighted by one student's remark: "[…] some students may have a better 'background,' such as [having sailed as] engine rating, than one who has never worked on a boat before." This commentary underscores the potential advantage conferred by real-world experience within testing environments that replicate practical scenarios. It is conceivable that an assessment conducted in such a milieu, where training and learning occur in alignment with the operational environment, might equip students with optimal competencies for their future professional roles. Consequently, a graded examination rooted in such an evaluative paradigm may yield superior grades for individuals more predisposed to excel in their initial work experiences.

Conversely, a student's assertion, "'This does not favour students without reading and writing difficulties'" raises pertinent considerations. This is particularly relevant in a program emphasizing practical orientation. Accommodating students with dyslexia aligns with collective interests, provided they exhibit commendable practical aptitudes.

Furthermore, the assertion that the assessment may afford preferential treatment to students favoured by instructors introduces an element of concern. While the practical assessment in question was ostensibly conducted

impartially, it is imperative to ensure that such equity is unequivocally perceived by students. This aspect warrants dedicated investigation as a potential area for further research. The goal is to fortify the integrity of practical assessments against perceptions of partiality or undue leniency.

The students' perspectives on assessment alignment to real world practice

In their seminal work, Säljö and Marton (1976) emphasize that students tend to tailor their learning strategies to optimize performance in summative assessments. Traditionally, the course in this experiment uses a supervised written exam without support materials besides a calculator. This favours students with strong theoretical knowledge and tends to focus on facts and superficial knowledge, which are easier to express in writing. Several studies have emphasized the importance of aligning assessment methods with learning objectives (Biggs, Thurner). The pursuit of good grades, crucial for securing desirable job opportunities post-university, often influences students to favour simpler exam formats such as multiple-choice over more complex free-response formats like essays (Van de Watering et al., 2008).

This experiment shows that students favour an assessment method that allows them to demonstrate proficiency in what they consider the most crucial learning objectives of the course. This preference is due to their understanding of the work tasks they will perform when they start working. This is interesting because it implies that students will be tested and graded in scenarios that closely resemble real-world situations they will encounter onboard a ship. This approach showcases the best practical students to future employers.

Sambell, McDowell, and Brown (1997) propose alternative assessments, such as practical simulator-based assessments, as more conducive to achieving quality learning outcomes. Students perceive these assessments as enabling, rather than hindering, the attainment of meaningful knowledge. The practical simulator-based assessment is particularly favoured in the current study, as it closely mirrors real-life scenarios and the competencies demanded in professional settings. This alignment with real-world requirements underscores the importance of collaboration between academia and industry in both teaching and assessment, as emphasised by Billett (2014).

In conclusion, the students' preference for practical assessments in this study is rooted in the belief that such assessments align more closely with the essential knowledge required for their future careers. The practical simulator-based assessment emerges as a valuable tool for evaluating competencies relevant to the intricacies of real-world scenarios. This necessitates ongoing collaboration between educational institutions and industry for effective teaching and assessment.

Can fairness in assessment compensate for the challenges with objective assessment?

Subjective and biased assessment by instructors is always a challenge, and the assessors' emotions play an important role (Gomez-Garibello & Young, 2018). In this study, there was no indication of biased assessment, and the instructors' use of assessment sheets was clear and concise. Nevertheless, students mentioned it as a possible challenge. The issue of unjust assistance is complex and will always be part of discussions among students. The assessment system must minimise its influence, but it is important to remember that some "assistance" given by instructors is meant to help students demonstrate their knowledge and thereby increase their scores. A student who can answer all questions without any assistance is a top student, but a student given "assistance" may also answer everything but will not be graded as highly. Clearly stating such rules to students before the assessment can reduce these suspicions.

To further address the suspicion of biased assessment, one idea is to allow students to follow the assessment process. At Unitec in Auckland, landscape architecture students could silently observe the assessment process of their exams. The conclusion was that this method positively affected the assessors' behaviour. It gave students a boost from the positive feedback comments about them and their work and fostered critical thinking among students. These interesting results can be applied to all assessments (Rennie, 2013).

The simulation used in this assessment requires two people to work together, one in the engine room and one in the engine control room, replicating an actual configuration on a ship. Since an engineer officer seldom works alone, this group assessment method mirrors real-world scenarios. They will seek, find, analyse, and fix errors as a team, using the entire team's competence

base. Research shows that students prefer individual assessment over group assessment, but they perform better in group assessments (Knight, 2004).

Although students were not directly asked about the group assessment, their comments showed positivity towards it. The interaction between students indicates the potential to broaden the scope of this type of assessment to cover Non-Technical Skills (NTS), an important part of the learning outcomes of maritime programs. According to the mandatory requirements for maritime education, as outlined in the International Maritime Organization (IMO) International Convention on Standards of Training, Certification, and Watchkeeping for Seafarers (STCW) (International Maritime Organization, 2010), students receive proficient training in these NTS. Assessing them in a practical-oriented summative assessment would be a bonus to the course and elevate this method to an even higher level of relevance.

Conclusion

This study has found that a practical simulator-based assessment method is preferred and well-received by a cohort of marine engineering students. They feel that this method accurately measures the skills and knowledge they need for their future careers. Given the practical orientation of marine engineering programs, summative assessments should also be as practical as possible to align with the learning outcomes.

The ability for students to verify that the assessment is truly unbiased, and objective is both interesting and important. Although no students reported feeling that the assessment was biased, several comments suggested that the possibility of bias should be investigated. Therefore, further studies should focus on the transparency of the assessment methods. The assessment method for this course will be further developed and tested in a new study when the course is next offered.

References

Billett, S. (2014). Integrating learning experiences across tertiary education and practice settings: A socio-personal account. *Educational Research Review, 12*, 1–13.

Cavanagh, A., Chen, X., Bathgate, M., Fredrick, J., Hanauer, D., & Graham, M. (2018). Trust, growth mindset, and student commitment to active learning in a college science course. *CBE-Life Sciences Education, 17*(1). https://doi.org/10.1187/cbe.17-06-0107

Gomez-Garibello, C., & Young, M. (2018). Emotions and assessment: Consideration for rater-based judgements of entrustment. *Medical Education, 52*(3).

International Maritime Organization. (2010). *International Convention on Standards of Training, Certification and Watchkeeping for Seafarers*. IMO.

Kibble, J. (2017). Best practices in summative assessment. *Advances in Physiology Education*, 110–119.

Knight, J. (2004). Comparison of student perception and performance in individual and group assessments in practical classes. *Journal of Geography in Higher Education, 28*(1), 63–81.

Kulamakan, K., & Rangachari, P. (2018). Beyond 'Formative': Assessment to Enrich Student Learning. *Advances in Physiology Education, 42*(1), 5–14.

Marton, F., & Säljö, R. (1976). On Qualitative Differences in Learning-II Outcome as a Function of the Learner's Conception of the Task. *British Journal of Educational Research, 46*(2), 115–127.

Oser, R., Gualtieri, J., Cannon-Bowser, J., & Salas, E. (1999). Training Team Problem Solving Skills: An Event-Based Approach. *Computers in Human Behaviour, 15*(3), 441–462.

Pérez-Sabater, C., Montero-Fleta, B., Pérez-Sabater, M., & Rising, B. (2011). Active Learning to Improve Long-Term Knowledge Retention. *Proceedings of the XII Simposio Internacional de Comunicación Social*, 75–79.

Rennie, J. (2013). Towards Assessment Transparency: Following on from 'The Crit', Can a Student's Learning Be Enhanced by Allowing Them to Witness Their Own Formative and Summative Assessment Event? *Critique Conference Proceedings*, 295–305.

Sambell, K., McDowell, L., & Brown, S. (1997). 'But Is It Fair?': An Exploratory Study of Student Perceptions of the Consequential Validity of Assessment. *Studies in Educational Evaluation, 23*(4), 349–371.

Sellberg, C. (2017). Simulators in Bridge Operations Training and Assessment: A Systematic Review and Qualitative Synthesis. *WMU Journal of Maritime Affairs, 16*, 247–263.

Struyven, K., Dochy, F., & Janssens, S. (2005). Students' Perceptions About Evaluation and Assessment in Higher Education: A Review. *Assessment & Evaluation in Higher Education, 30*(4), 325–341.

Van de Watering, G., Gijbels, D., Dochy, F., & Van der Rijt, J. (2008). Students' assessment preferences, perceptions. *Higher Education, 56*, 645–658.

Appendix 7.1

In the first part of the checkout. you were to fire up the man engine for the sea voyage. The evaluation form was a checklist with the grading performed / passed / not passed. There were 33 items on the list, and 29 items had to be passed for the test to be approved. Additionally, 5 of them were considered critical and all of these had to be approved.

Questions regarding scenario 1					
1. The check list is a good way to show your knowledge	1 Strongly agree	2 Agree	3 Neutral	4 Disagree	5 Strongly disagree
2. I mean this method ensures equal opportunities for the students	1 Strongly agree	2 Agree	3 Neutral	4 Disagree	5 Strongly disagree
3. I mean this method reduces the possibility for biased treatment of the students	1 Strongly agree	2 Agree	3 Neutral	4 Disagree	5 Strongly disagree
4. I mean this method is transparent with the possibility to verify the assessment	1 Strongly agree	2 Agree	3 Neutral	4 Disagree	5 Strongly disagree
Additional comments					