ARTICLE

Integrating motivated goal achievement in maritime simulator training

Simen Hjellvik[1] · **Steven Mallam[1]**

Abstract

The application of cloud-based simulators has increasing momentum in maritime education and training as a virtual platform for supplementing professional training with task-specific simulation-based learning. Congruent with traditional simulator technology and training methods, this new condition allows for asynchronous and unlimited simulator access where participation in training sessions is at the discretion of the trainee. Furthermore, this provides a learning environment that can have adaptive features to the trainee characteristics with selectable complexity levels, automated feedback, and automated performance assessment. With the advent of this emerging training technology, the organization of how training is deployed, the new roles and expectations for trainees and instructors, and the application of new data-driven analytics are not yet well understood. This paper employs a quasi-experimental study to investigate trainee *motivation*, *personality traits*, and *task performance* using a novel training design. The study was administered remotely to a sample ($N = 18$) of first-year maritime students who first completed a knowledge acquisition phase before conducting a skill acquisition phase with repeated simulator training of a procedural task, and finally a performance assessment of the simulator scenario. Data was collected from (1) a multiple-choice knowledge test, (2) a short-scaled Big Five Inventory, (3) the Motivated Strategies for Learning Questionnaire, and (4) programmed simulator metrics. Results indicate that *Extraversion* correlates with *motivation* and *self-efficacy* predicts training performance. No significant relationship between prior knowledge and the training performance was found. The results can provide knowledge for implementation and delivery of remote simulator training in maritime education and training, as well as other fields.

✉ Simen Hjellvik
Simen.Hjellvik@usn.no

1 Department of Maritime Operations, University of South-Eastern Norway, Campus Vestfold, Raveien 215, 3184 Horten, Norway

1 Introduction

Immersive technologies for transforming knowledge into skills are gaining momentum in maritime education and training (MET). MET refers to the field of education and training employed by institutions that provide theoretical and practical development of professional maritime competence. MET utilizes different strategies and technologies in its learning processes and continually evolves in order to maintain relevance in an ever-changing industry (Sharma & Nazir, 2021).

There is no single definition of simulation; however, the important features of modern simulators are the presentation of realistic situations that engage the trainee in activity that provides practice opportunities on meaningful tasks (Lajoie, 2021). In the concurrent context, the definition concerns computer-based simulators which apply a range of immersive features to the user, from allocentric 2D and 3D virtual reality desktop simulators to fully immersive full mission simulators and egocentric head-mounted virtual reality (DNV, 2021). Traditional maritime simulators have typically been limited to desktop devices and full mission simulators physically located and fixed within a training facility (Mallam et al., 2019), whereas novel cloud-based simulators (CBS) and immersive technologies pave way for a more distributed paradigm. In their meta-analysis on simulation-based learning in higher education, Chernikova et al. (2020, p. 500) state " … little is known about for whom simulations are particularly helpful, what scenarios are effective, and what additional supportive structures makes them effective for learners with different learning prerequisites." Thus, even as simulation technologies advance, enabling new and different learning opportunities, basic pedagogy-related questions with regards to simulation deployment continue to remain unclear.

CBS is a back-end technology where remote servers compute the whole experience which is delivered to the user's device through a high-speed network connection, rendering a lower hardware requirement for the user's device (Motejlek & Alpay, 2021). CBS technology can create a virtual learning environment that has the same interaction features of E-learning, but with heavier computation-demanding content as found in traditional on-campus simulator training. Furthermore, synchronous or asynchronous training sessions can be organized between peers and instructors, or individually and on-demand. This reduces trainee barriers to simulator access associated with brick and mortar training institutions, theoretically enabling unlimited training opportunities at the trainee's discretion. Current CBS applications create new opportunities for how training is organized and delivered to trainees, including the social structures and implications. Remote training delivery with programmed system automation makes CBS a good candidate for creating adaptive training, where instruction is manipulated to provide the trainee with an effective learning experience (Landsberg et al., 2012).

Various facets of motivation can be a precursor, a mediator, or a concomitant outcome of the learning process (Zimmerman & Schunk, 2012, pp. 1–30). With what is known, sources of motivation can be leveraged when designing a training delivery to create an effective learning environment. Viewing the trainee as

an individual creates a challenge to provide feedback at the level of the trainees' individual characteristics, and more so when the feedback is given through training that is standardized and situation specific. Organizing feedback as automated pre-programmed responses of a simulator could include the concepts of *feed-up*, *-back*, and *-forward* (Hattie & Timperley, 2016). An assessment should follow the completion of any learning or training effort to identify the degree to which the achievement approximated the goal and how relative expectations prove to be.

This study investigates traditional marine machinery simulation, transcended to a new delivery method by the back-end technology of cloud-based simulation. This study contributes with an exploration of motivation and personality traits that indicates the importance of training design to facilitate progress in the learning process.

This study focuses on designing and conducting marine machinery simulation in an integrated manner with the relevant education courses through the phases (1) lecture-based knowledge acquisition, (2) asynchronous skill acquisition, and (3) performance assessment, as will be explained later. The study was integrated in marine machinery courses for first-year marine engineering students at one college and one university, and in marine engineering courses for first-year nautical sciences students at one college and one university. Measurements of personality traits and motivation are collected from the trainees, and metrics from the simulator are collected for training performance indicators and training quantifications that contribute with insight into progress patterns.

In this study we consider the learning process to cover the complete course integration, i.e. from the lectures to the end of the simulator training and assessment. Integration of the study into the relevant marine machinery courses means that the learning process covers learning objectives already existing in the course descriptions: (1) knowledge of steering gear systems, (2) knowledge of steering gear operation, and (3) experience of steering gear operation by simulator training. By such integration, the learning objectives were delivered to all students in these courses on behalf of their course responsible lecturers. Training process refers to the measurements generated by the trainees' repeated simulator training and was used to create quantifications. These training quantifications can help the instructors interpret the trainee's intervention beyond traditional assessment of performance metrics. Training or test performance indicators from the *Training scenario* and the *Test scenario* refer to the automated assessment scores, namely *Best training score* and *Test score*. The relation between these measurements is addressed as shown in Table 1 and tested accordingly.

2 Background

2.1 Motivation and personality

Motivation is the internal process and causal stimulus for actions and behaviour from intrinsic and extrinsic factors (Schrader et al., 2021). In the context of training, motivation refers to the direction, intensity, and persistence of learning-oriented behaviour (Colquitt et al., 2000) before, during, and after the activity (Gully &

Table 1 Research questions and hypotheses

	Statement	Operation
RQ1	How is motivation and personality related in this learning process?	
H1	*Personality trait factors correlate with motivation scales.*	Correlation
H2	*Personality trait factors are facilitators of motivation.*	Regression
RQ2	Is test performance a product of training performance or training process?	
H3	*Training performance will be different than test performance.*	Related samples
H4	*Test performance will be predicted by the training process.*	Regression
RQ3	Is motivation related to the training process and training performance, or test performance?	
H5	*Motivation scales correlate with the simulator metrics.*	Correlation
H6	*Training performance will be predicted by motivation.*	Regression

Chen, 2010). Proactive qualities like personal initiative, perseverance, and adaptive skill emerge from motivational beliefs and self-regulatory learning strategies which are reciprocally interactive (Zimmerman & Schunk, 2012). In online technology-based training proactivity is an important factor in relation to planning time and effort dedicated to the training program and to overcome the technical difficulties and learning process interruptions usually occurring with such training delivery (Bell et al., 2017).

Motivation and cognition are key factors to develop self-regulatory learning strategies for academic achievement. Motivation is a factor that is beneficial from an early stage as the will to learn, whereas metacognitive strategies stem from practice and instruction as the skill to learn (Zimmerman, 2008). Early-stage learners would be more perceptible to motivational stimulation from the learning environment than benefiting from strategies yet to be developed. Also, motivational beliefs and feelings are found to be greater with experts than with non-experts and novices; however, the motivation of novices is greater with facilitation than without (Zimmerman, 2008).

People can hold on to motivational beliefs that are based on incorrect knowledge, even after being presented with correct explanations (Pajares, 2012). Following this, it is the newly acquired beliefs that are most prone to be changed. Motivational change during training is usually a secondary outcome where skill-acquisition based on the prior knowledge foundation is the primary outcome. Altogether, training is usually designed to target declarative knowledge and skill-acquisition; however, motivational outcomes should be included to provide a more complete profile of the learning process (Kraiger et al., 1993).

Cognitive ability is an important predictor for individual task performance and positively affected by motivation to learn (Kanfer & Ackerman, 1989). However, beyond individual capabilities, trainee characteristics influencing the learning process also include personality, motivational constructs, values and interest, attitudes and emotions, and perceptions (Bell et al., 2017; Gully & Chen, 2010). Declarative knowledge, skill-acquisition, and self-efficacy as training outcomes are found to be predicted by both cognitive ability and motivation to learn, thus proving the

outcomes of training to have causation beyond cognitive ability (Colquitt et al., 2000). With digital learning environments management of mental load is a core principle to ensure engagement with the learning program where optimal germane mental load is important for effective learning to occur (Clark, 2021).

Moving on to the five-factor model, which was developed as five abstract factors of personality traits, namely *Extraversion*, *Agreeableness*, *Conscientiousness*, *Emotional Stability*, and *Openness to Experience*, summarizes numerous specific personality characteristics and creates a profile (John & Srivastava, 1999). The personality model described by these dimensions represents a variety of traits across the individual. The profile can be subject research and where in particular the individual dimensions are of interest. For example, *Extraversion* is found to have a strong correlation to positive emotional dispositions (Shiota et al., 2006) and academic motivation (De Feyter et al., 2012), *Conscientiousness* is found to be a strong predictor of academic performance (Barrick et al., 2001; De Feyter et al., 2012), and motivation is proposed to mediate the relationship between personality and performance (Parks & Guay, 2009).

To sum up, the variety in individual trainee characteristics is important for designing training programmes. Decisions must be made whether designing training programs for (1) a selected population, i.e. trainees with identified traits that positively predict a successful performance (Towler & Mitchell, 2014); designing to (2) mitigate individual differences, i.e. target a broad population and lower or adjust the goal; or (3) designing training programs that adapt to and support the population variance towards the same standardized goal.

Furthermore, investigating universal constructs that hold predictive premonition for learning outcomes, like motivation and personality as discussed above, gives more insight into the individual than assessing the correlation between training performance and previous academic performance.

2.2 Feedback and assessment

Feedback provides trainees with information about their performance or behaviour. It can give an evaluation of their understanding or performance which identifies error, helps repair faulty knowledge, and gives cues to change strategy (Johnson & Marraffino, 2021). If treated as a unitary variable, feedback is any process that gives the learner an instructional response on a continuum from correct or incorrect to substantial corrective information (Kulhavy, 1977). Along this continuum the complexity advances from the binary format to providing more information up to a point where the feedback becomes a new instruction (see Fig. 1) (Hattie & Timperley, 2016; Kulhavy, 1977). The goal of instruction is to motivate the trainees at a sufficient level of intensity to sustain an effort to make sense of the material presented (Mayer, 2021).

For the reason of filling the gap between comprehension and what should be known, feedback must be applied to a context where it can add to an existing knowledge base. In active learning the trainee makes personal choices as what to learn and what to leave behind (Bell et al., 2017). Subsequently, the effectiveness
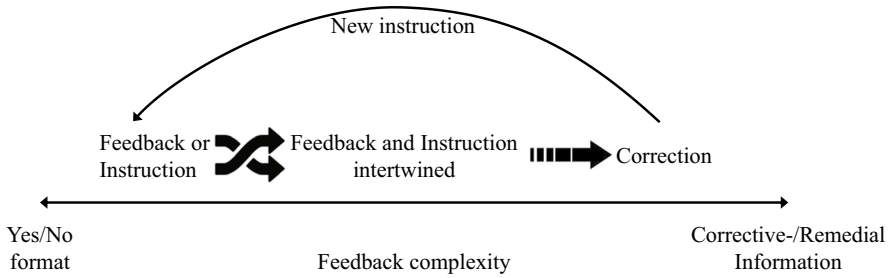
Fig. 1 Feedback complexity model (Hattie & Timperley, 2016; Kulhavy, 1977)

of feedback on performance thus depends on the recipient accepting, modifying, or rejecting it as abstruse or comprehendible information (Kulhavy, 1977). In view if this, feedback should include the three facets *feed* (1) -*up*, i.e. clarify the goal; (2) -*back*, i.e. assess present performance; and (3) -*forward*, i.e. guide consecutive actions and be strategically focused on the correct level (Hattie & Timperley, 2016). Targeting different levels of the learning process, feedback at the *task* level would include corrective feedback on results to distinguish correct from incorrect performance. Its aim is to enhance the knowledge base with more or different information and is more powerful when addressing faulty interpretation than the lack of knowledge. Feedback at the *process* level would be specific to the underlying and related processes of the task, such as the effect of the task prosecution on the goal and related system processes. Beyond acquisition, storing and applying knowledge, *process* feedback will address the trainee's strategies for problem-solving based on error detection and identifying knowledge gaps themselves.

Turning to assessment, this is the activity that provides the information for feedback at the task, process, or self-regulation level, i.e. information that discloses the present state and the discrepancy towards the goal. As such, an assessment can include all the levels of feedback whilst evaluating current proficiency. For assessment feedback to be effective the information collected should identify the performance relative to a specified and understood goal (Hattie & Timperley, 2016). Likewise, a feature of such measures is that they can be used in a training delivery with repeated sessions to change the trainee's self-efficacy as the confidence in one's ability is calibrated against actual performance (Zimmerman, 2008). Such calibration through repetition should strengthen self-judgement, which fundamentally hold the inductive risk of overconfidence in one's abilities (Dunning, 2018). In summative assessment after the knowledge and skill acquisition, tests evaluate the accuracy and validity of performance or retaining knowledge as the crystallized product of previously learned content (Lindner, 2021). In addition, assessment through performance in a virtual environment will also measure the mental model construction of previous learning (Lindner, 2021), and likewise, training and being tested in the same environment should avoid extraneous cognitive demand.

2.3 Simulator technology

As explored below, CBS technology adds new features to existing technology. Kim et al. (2021) analyse the technical properties of maritime CBS and find the lack of social interaction and the lack of formative assessment to be its major weaknesses. In the context of traditional on-campus simulator training, these two factors are interlinked as contingent on the physical presence of the instructor and peers during the session. With CBS technology these two concerns are in fact technology dependent and this study will explore a design that to some extent provide formative assessment through programming of the scenario. To counter the assumed deficit of no physical peer and instructor interaction, the effectiveness of the training delivery is more dependent on design and automated feedback to provide the targeted learning outcomes. As a result, the instructor functions as a facilitator rather than a purveyor of knowledge. What is more, an autonomy-supportive role of the instructor has been found positively related with trainee engagement, motivation, and performance (Towler & Mitchell, 2014).

The CBS platform in this study is designed for creating the decentralized and remote delivery of a learning environment that is administered synchronously or asynchronously by the instructor. To elaborate, the interaction of asynchronous training is a student-centred application that allows different learning strategies (e.g. on-demand and flipped classroom) (Gurtner, 2014) than an instructor-driven synchronous learning interaction. For instance, Kinshuk (2016, pp. 111–124) warns about the possible misunderstandings in synchronous collaboration with digital learning environments as the social condition for communication is different than the face-to-face setting it is substituting. Instead, with this CBS technology there is the opportunity to create a virtual instructional agent that interacts at a one-on-one level with the trainee. To review, one-on-one instruction has been shown to be vastly more effective than collective instruction (Bloom, 1984). Approaching the trainee individually through a structure with adaptive and personalized features can be a feasible approach to investigate the use of CBS in an integrated learning process that is not possible with the traditional on-campus simulator training. Therefore, the ubiquitous availability of CBS creates new opportunities to integrate simulator training with classroom learning outcomes that previously were inconvenient or ineffective to administer. In contrast to traditional on-campus simulators, CBS training can be delivered to the trainee whenever and wherever to prepare for the next lecture and on-campus simulator training or repeat training exercises until knowledge and skills crystalize. Consequently, this could liberate more on-campus time for lecturing or enable better quality of the on-campus simulator training as the trainees are given a take-home simulator to use in their studies. Turning to the instructor, in the maritime context the instructor of simulator training in education is required to have competence according to maritime standards on simulator instruction. The instructor is expected to determine the suitability of exercises to meet learning objectives and also the suitability and behaviour realism of the simulators for the learning objectives and training exercises (DNV, 2022). In view of this, implementing CBS would induce different tasks on the instructor as will be explained later.

Moving on to interactional features, behavioural realism is a technical consideration of the degree the simulator's virtual environment resembles real equipment (DNV, 2021). The interaction with the user and how well this environment is projected to the user can be defined through the immersive level of the technology (Makransky, 2021). Immersion can be viewed as the level of involvement experienced in a virtual environment which disconnects the user from the real surroundings (Radianti et al., 2020). Further, immersion has been found to be positively related with enhanced learning processes, student engagement, and learning outcomes, as well as experiential concepts like presence and flow (Suh & Prophet, 2018). And yet, immersive virtual environments do not necessarily improve learning on its own; it is the instructional methods applied that improve learning (Makransky, 2021). To review, the CBS of the study is classified as a virtual reality simulator (DNV, 2021), and the technological view of immersion categorizes CBS as lower immersion due to the use of desktop monitors, regardless of the perceived user experience (Parong, 2021). The simulator's fidelity is high with a detailed system realism, and any level of experienced presence facilitates engagement with the learning environment which adds value to the learning experience. In sum, immersion is a feature that can facilitate increased perceived presence, and research finds presence to correlate with motivation, engagement, and positive emotions. Nevertheless, presence alone have not shown effects of increased learning (Feldon et al., 2021).

2.4 Adaptive and personalized learning

In the following section, user-centred design can take the approach of adaptive and personalized learning systems. Kinshuk (2016, pp. 165–176) defined smart learning environments as ecosystems where technology and pedagogy inter-fuse in the individual's learning process. Here the trainee would be able to seamlessly transfer knowledge between learning environment contexts (Kinshuk, 2016, pp. 29–40), e.g. from self-studying to classroom lectures, then to on-campus simulator training with peers, followed by individual CBS repetition and more self-studying. A strength of the CBS platform in this study is that training scenarios can be designed by programming to incorporate features of smart learning systems (Tabuenca et al., 2021), which (1) sense the behaviour pattern and performance of the trainee's interaction, (2) analyse the data detected or collected from the system manipulation of the interaction, and (3) react to the interaction with direct feedback, system variable adaptions, and performance assessment. As explored in this study, CBS platforms of these features could integrate with current educational practice towards features of a smart learning environment.

Adaption of the learning process is the system's response to user interaction where actions are triggered to change system parameters which create a customized environment that engages the user (Vesin et al., 2018). The intention of an adaptive training system is to provide a closed-loop feedback system (see Fig. 2) that can change as a response to how well the trainee is performing (Kelley, 1969). In real-time or in repetition of a training task there needs to be one or more adaptive variables that adjust the complexity by the training system
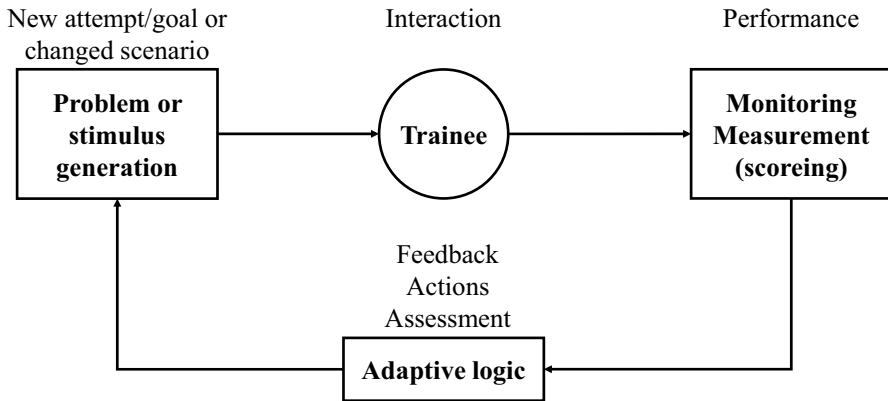
New attempt/goal or   Interaction   Performance
changed scenario

Fig. 2 Closed-loop adaptive training adopted from Kelley (1969)

automation. As discussed previously, individual trainee characteristics can moderate the effectiveness of feedback; thus, an adaptive training system is a natural progression to mitigate this in multimedia learning (Johnson & Marraffino, 2021). To personalize the use of the simulator environment the trainee decides the amount, pace, and difficulty of training to achieve their preferred level of mastery (Vesin et al., 2018). Personalization of learning environments can increase the trainee's satisfaction as the content is tailored to the trainee's performance level and information overload is attenuated (O'Donnell et al., 2015). In brief, this level of individual personalization, i.e. remote on-demand training, is the key feature that discriminates CBS from traditional on-campus simulator training where the same training strategy is given an enhanced application.

For simulator training to be effective trainees must be presented an environment with fidelity (Grossman et al., 2014) that provides pedagogical opportunities to create substantial changes in individual knowledge, skills, or attitudes that are needed for the competencies of the job in training (Salas et al., 2012). The CBS platform provides a high-fidelity virtual environment that is capable of facilitating a training strategy that gives (1) *instruction* of the knowledge trainees need to learn; (2) *demonstration* of the knowledge, skills, or attitudes expected; (3) *practice* of these; and (4) *feedback* for remediation of performance and learning (Salas et al., 2012). In summary, this study applies the CBS features of (1) no physical and social interaction, (2) virtual instructional agents, (3) remotely operated, (4) delivered asynchronously from the learning program, and (5) smart learning environment capable with adaptive and personalized features. Although the CBS holds the same low-immersion/high-fidelity feature as its antecedent technology, the traditional on-campus desktop simulator training is physical, social, and synchronous. This positions CBS to serve a learning program with mass repeated training that otherwise would not be possible.

3 Method

3.1 Experimental design

The study was integrated as a substitute for the relevant learning outcomes in marine machinery courses at 4 different institutions. All phases of the whole learning process as described in Sect. 1, including the experiment, was remotely disseminated through the trainees' personal computers. As such, no physical on-campus desktop simulation could be set up for a technology comparison control group. As the trainees' interaction with the training program was under investigation, a population control group (e.g. senior student or expert group) would also be unfeasible as this would only compare previous cohorts' uncontrolled and retaining competence against the present sample. A no-treatment control group, i.e. *Test scenario* only, from the same sample population was also decided against as random assignment of such would not achieve to evenly mitigate the demographic factors of discipline and affiliation. A within-group quasi-experimental design was favoured to conserve the statistical power of the recruited sample. Resulting from this, all recruited trainees were given the same treatment and instructions.

3.2 Instruments and experiment variables

A 10-item *Knowledge test* was created as an online survey to be administered between the knowledge acquisition and skill acquisition phase. The content of the test reflected the lectures given and the intended learning outcome.

A demographic questionnaire was administered as an online survey to collect *age, gender, institution affiliation*, and relevant *maritime work experience*.

A Big Five Inventory (BFI) was collected through an online survey. The Norwegian-translated and shortened 20-item version (Engvik & Clausen, 2011) of the BFI-44 (John & Srivastava, 1999) was used on a 7-point Likert scale. The tool provides a reliable indication rather than the full profile of the facets within the personality traits *Extraversion*, *Agreeableness*, *Conscientiousness*, *Emotional Stability*, and *Openness to Experience*.

The Norwegian translation (Kvinge & Engelsen, 2016) of the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich & et al., 1991) was collected on a 7-point scale through an online survey after the experiment. It is a retrospective self-report instrument developed to measure academic students' motivational orientation and their different use of learning strategies. The 31-item motivation section comprises six subscales that focus on the students' goals and value beliefs for a learning task, beliefs about their ability to succeed in the task, and their anxiety for being assessed in the task (Duncan et al., 2015):

- *Intrinsic Goal Orientation.* The degree to which participating in the task holds a value of its own, such as challenge, curiosity, or mastery. High response

indicates the perception that participating in the task is an end in itself, rather than a means to an end.

- *Extrinsic Goal Orientation.* The degree to which participating in the task is valued by external factors such as grades, rewards, performance, competition, or evaluation by others. High response indicates the perception of participation as the means to an end, where the participation in itself is not of greater value.
- *Task Value.* The evaluation of how important, interesting, and useful the task is. High task value should lead to greater involvement in the learning task as result of the perception of the importance, interest, and utility of the task content.
- *Control of Learning Beliefs.* The belief that effort to learn will result in desired outcomes. Perception of one's own academic performance should lead to provide the strategic effort needed to effect the desired changes.
- *Self-Efficacy for Learning and Performance.* The expectancy for performance success and the self-appraisal of one's ability to master the task. High self-efficacy indicates judgements about one's abilities and confidence in one's performance.
- *Test Anxiety.* Probes both the cognitive component and the emotional component of anxiety. The negative worry that disrupts performance and the affective and physiological aspect of anxiety.

All BFI 20-N factors and all MSLQ motivation scales, except *Task Value*, were found to be normally distributed according to the Shapiro-Wilk test. Alpha reliability of the data is reported in Table 2 with comparison to the original studies and the MSLQ reliability generalization study of Holland et al. (2018).

Table 2 Alpha reliabilities of data and related studies

Scale/factor	N items	α	Engvik and Clausen (2011)	Kvinge and Engelsen (2016)	Holland et al. (2018)	Pintrich and et al. (1991)
Extraversion	4	0.811	0.78			
Agreeableness	4	0.457	0.63			
Conscientiousness	4	0.689	0.57			
Emotional stability	4	0.592	0.73			
Openness to experience	4	0.382	0.63			
IGO	4	0.654		0.78	0.709	0.74
EGO	4	0.296		0.75	0.692	0.62
TV	6	0.788		0.82	0.833	0.90
CLB	4	0.710		0.67	0.645	0.68
SELP	8	0.873		0.86	0.879	0.93
TA	5	0.539		0.73	0.759	0.80
M_{POOLED}	31	0.854				

$α$ = Cronbach's alpha of dataset, IGO = Intrinsic Goal Orientation, EGO = Extrinsic Goal Orientation, TV = Task Value, CLB = Control of Learning Beliefs, SELP = Self-Efficacy for Learning and Performance, TA = Test Anxiety, M_{POOLED} = Pooled Motivation Scales

Table 3 Collected simulator performance indicators and training quantifications

Variable	Type	Description	Unit
Total training attempts	TPI	Total attempts at the Training scenario, excluding the final Test scenario attempt.	Number
Total training time	TPI	Total accumulated time spent during the training attempts.	Seconds
Best training scenario score	TPI	Best score from all Training scenario attempts, usually the last.	Scale
Test scenario score	API	Score of the final Test scenario attempt.	Scale
Time Demand	TQ	$TD = \frac{Score\ range\ slope+intecept}{60\ min}$	Minutes
Time Retention	TQ	$TD = TD - \frac{Score\ range\ slope+intecept}{60\ min}$	Minutes
Progress Rate	TQ	$PR = \frac{Score\ range}{TD}$	Score/minute

TPI = training performance indicator, API = assessment performance indicator, TQ = training quantification, TD = Time Demand, TR = Time Retention, PR = Progress Rate

Simulator metrics were automatically collected as programmed in the scenario exercises. The collected metrics was used for performance indicators and quantifications (see Table 3). Training performance indicators refer to the performance metrics derived from procedure training in the *Training scenario*. Assessment performance indicators refer to the performance output from the *Test scenario*. The quantifications were made to help describe the process of repeated training and the state of proficiency at the point when training was concluded. They were made from the linear regression of all *Training scenario* attempts by each trainee using their scores and accumulated time dissipation, as such they can be used as the training process is ongoing to indicate how it is going. In short, *Time Demand* is the total time the trainee needs to spend training in order to achieve a full score. *Time Retention* is the amount of time the left to train in order to achieve a full score. *Progress Rate* is the pace in which the trainee is able to acquire new score point towards a full score.

3.3 Sample

The sample ($N = 18$) was recruited from first-year students enrolled in a Nautical Sciences ($N = 12$) and Marine Engineer ($N = 6$) programme at university and college levels in Norway. The average *Age* was 23.7 years ($SD = 5.39$), and relevant *Maritime Work Experience* was on average 2.9 years ($SD = 3.31$). *Age* and *Maritime Work Experience* was found to be non-normally distributed according to the Shapiro-Wilk test. Data was only collected after written informed consent according to the filed Norwegian Centre for Research Data notification (no. 753508). The Norwegian Universities and Colleges Admission Service (2020) inform that a total of 588 students started in these education programs. The Database for Statistics on Higher Education (2020) indicates a 9.2% dropout across all maritime studies before this study was conducted. Consequently, the sample represents a 3.4% extraction of the national cohort.

3.4 Simulator model

This study uses the K-Sim Engine simulator platform which can operate through remote cloud services or locally installed on computers. It supports several models that replicate different types of maritime vessels and machinery with both 2D desktop interface and 3D desktop VR as shown in Fig. 3.

The user interacts with the simulator through one or more desktop monitors, keyboard, and mouse. The model used in this study was the MAN 6S70ME-C SCC, which replicates the machinery system of a 152.000 DWT Suez Max oil tanker.

The CBS training is performed with the 2D interface (Fig. 3, left) on the trainees' personal desktop computers. The lectures (Fig. 6) used both the 2D model and the 3D desktop VR model (Fig. 3, right). The 3D desktop VR model is an allocentric (i.e. looking into the environment through a desktop monitor) virtual reality model that was currently not available for egocentric interaction (e.g. through head-mounted devices). This virtual reality environment had a 360-degree of freedom first-person view, and the same interactive simulator model as with the 2D simulator model of the CBS.

Visual animations indicate system functioning, such as visualization of valves which are taken from a closed position to an interim position before it becomes fully open. Each single machinery system component (e.g. valve, valve actuation motor, or actuator control automation) is built up of several variables that enable its dependent system function and to connect with other components or functions of the simulator model. To replicate the complexity of a real-life machinery system the simulator variable is default-programmed to behave accordingly. This web of interacting variables provides the simulator model with the expected natural response to all interactions from the trainee. However, all default-programmed variables can be controlled by force of a pre-programmed scenario, such as with a training exercise.

3.5 Scenario programming and exercise

For a training scenario such as this to be effective it needs to be systematically approached (Grossman et al., 2014) by (1) identifying the trainee's knowledge and skill's inventory before training; (2) identifying task or competences targeted with the training; (3) which specific and measurable learning outcomes (Kraiger et al.,
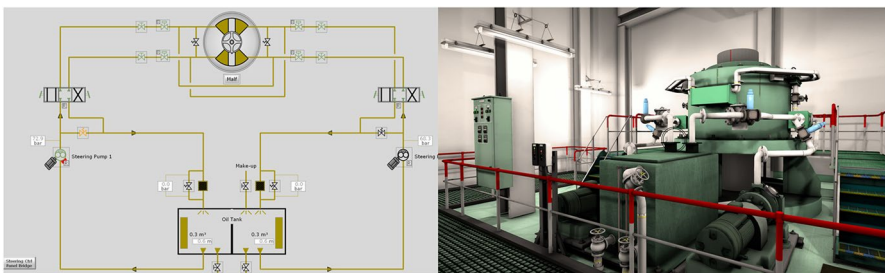

Fig. 3 K-Sim Engine 2D interface and 3D desktop VR interface

1993) to be facilitated by the design of the training; (4) designing a training scenario that triggers the demonstration of the target competences; (5) incorporating performance measures that assess the learning outcomes and set standardized levels of these performances; (6) the performance of the trainees is diagnosed accordingly, during or after the training scenario; and (7) the performance diagnosis is used for feedback to the trainee.

The scenario selected for the study was for trainees to perform a safety critical pre-departure procedure on the vessel's steering gear machinery and remote controls. The steering gear is one of the vessel's most critical systems during the voyage as it directly controls the rudder. The procedure includes visual control and functional tests with the steering gear system and its components, then tests of its local and remote operation from the steering gear room and the bridge, respectively. The trainee was instructed to assume and perform the tasks of both a navigational officer on the bridge control station and the tasks of an engineer officer on the equipment in the machinery area.

Programming scenarios with this simulator model are typically performed by or in collaboration with the capacities of (1) a subject matter expert, (2) an experienced simulator instructor, and (3) advanced familiarity with the simulator model and programming software. A challenge with discipline-specific learning environments such as this is that the scenario author rarely holds the subject matter, pedagogical, and technical competence necessary to author the scenario alone as done in this study (O'Donnell et al., 2015). Programming the exercise in this study was performed as illustrated in Fig. 4, and consists of building forced system behaviour
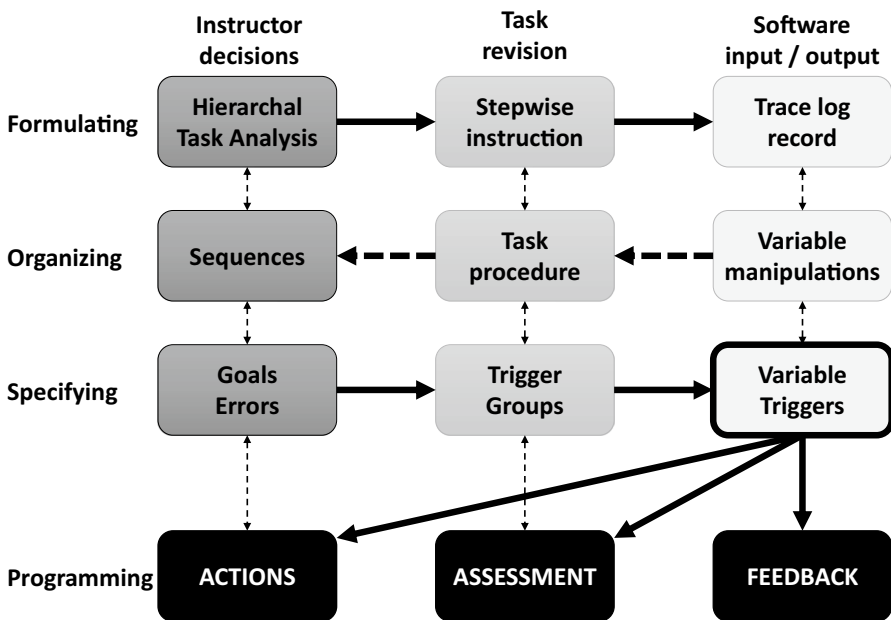

Fig. 4 Back-end programming approach with the Neptune Instructor software for K-Sim Engine models
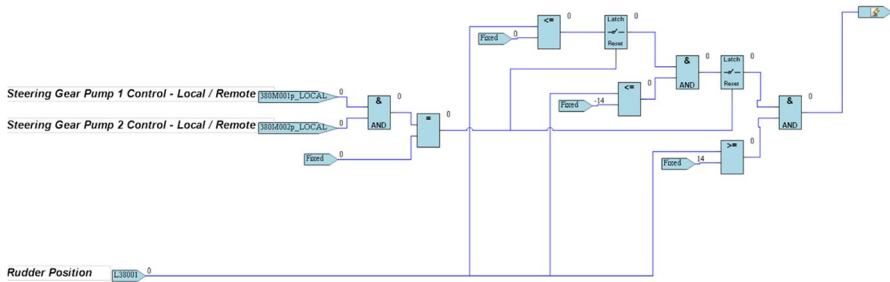
Fig. 5 Variable Trigger that detects a specific interaction with manipulation of simulator variables in a particular sequence

Table 4 Selectable conditions of the scenario

Scenario level	Objective	Condition
Test	Performance assessment	Summative assessment only, no supportive structures.
Training	Procedure training	Assessment as formative feedback. Instructional correction, reminders, feedback, and feed-forward as supportive structures.
Explore	Failsafe free play	No assessment or supportive structures.
Information	Semantic system knowledge	System and component descriptive information only.

Actions, collecting and communicating *Assessment* metrics, and providing automated *Feedback*. The feedback can typically be in the form of text from virtual instructional agents, multimedia content, or internet resources.

The purpose of the approach outlined in Fig. 4 is to create *Triggers* from the simulator variables that are used to build the exercise. As illustrated in Fig. 5, the triggers are built from simulator variables and logic blocks to give a binary output activated by a particular interaction or behaviour from the trainee.

Creating the adaptive environment was performed by designing different complexity levels of the scenario programmed according to Table 4. When starting the exercise, a choice of these 4 scenario levels was presented to trainees. These selectable alternatives each governed the activation of their separate set of *actions*, *assessment*, and *feedback* with the deactivation of the other sets, resulting in 4 independent conditions of the exercise scenario. In the *Training scenario* and *Test scenario*, adaptive features to the system behaviour were added with *actions* triggered by user interaction. Adaptive *feedback* was programmed for the *Training scenario* and the *Information scenario*. Performance *assessment* was triggered and given at the conclusion of the *Training scenario* and the *Test scenario*. For instance, in the *Training scenario* the trainee could try to operate a machinery system that was without electrical power, and if doing so, an instructional agent would appear making the trainee aware of this error, reminding to check electrical power supply before operating such systems, and guiding the trainee towards where and how to restore an operable condition of the system. Further, instructional agents could appear and provide cues on important parameters to be aware of once the trainee would start to

manipulate a machinery system. Simultaneously, *action* triggers would be activated to make the premonition occur sometime during the scenario, if not being tended to as prescribed. Throughout the course of performing the procedure, the instructional agents would confirm the success of completing steps and remind what to start with next. On the contrary, in the Test scenario no supportive structures in form of instructional agents would interact with the trainee. However, the procedure and the performance assessment were the same in both these scenario levels.

The CBS platform with the simulator model and the scenario exercise was available to the trainees for the duration of the spring school semester (approximately 4 months in total) 2021. Records of all their activities on the platform were available for the trainees during their access period, including the automated performance assessment from any scenario attempts. One issue that occurred during the programming phase was with the capacity of the simulator model which would start to lag if running a highly complex programmed scenario. There was no issue with the amount of content or text, but triggers needed to be programmed in the simplest layout as possible for its function.

3.6 Procedure

First, a knowledge acquisition phase was administered through online video conference with two 45-min lectures on steering gear systems and pre-departure procedures. Then a third 30-min lecture was distributed as a pre-recorded 3D VR multimedia video. The multimedia video lecture was a screen recording of the lecturer performing the procedure task and simultaneously elaboratively explaining the procedure in a 3D virtual reality environment. After the lectures, the trainees inscribed the online *knowledge test*, informed consent forms, demographic survey, and BFI survey.

Second, the skill-acquisition phase consisted of simulator training with the CBS training on the procedure task. The training delivery was asynchronous and thus deprived of all social supportive structures such as instructor and peer interaction. The trainees were given access to the CBS platform and instructed to focus on repeated attempts with the *Training scenario* level and revisit the *Information* and *Explore scenario* levels to address any emerging knowledge gaps. No instruction was given to what level of performance was considered a sufficient task performance of the *Training scenario* level. The trainees were instructed to self-assess their performance according to the 220-point maximum assessment score automatically given by the simulator after each attempt. The CBS automatically stored user activity and the assessments for collection.

Third, once satisfied with their performance on the *Training scenario*, the trainees would proceed to the *Test scenario* level. This level was designed as a pure examination with no supportive structures such as included in the Training scenario level, as described in the previous section. Shortly after the trainees completed the final scenario and was given their final performance assessment, the assessment phase was concluded by directing the trainees to the online survey containing the MSLQ.

Finally, after completing all phases (see Fig. 6) the trainee was discharged. Then, raw data from the online questionnaire service and CBS platform was collected and anonymized. The original recordings were then deleted from third-party services.

3.7 Data analysis

The raw data was screened for non-response on the surveys and inactivity during the CBS training. One case of non-response on the BFI questionnaire was detected and two cases of inactivity on *Training scenario* attempts were identified. These data points were discarded after verifying with the participants. The BFI and MSLQ items were accumulated to composite scores according to instruction by their respective manuals (Engvik & Clausen, 2011; Kvinge & Engelsen, 2016). The collected data was analysed with SPSS 26.0 before exported and visualized as presented in the paper.

4 Results

Table 5 summarizes the collected data from the study. The trainees participating used the CBS training platform over multiple attempts, with days or weeks between their sessions. Some preferred shorter time dissipation and more attempts, whilst others spent more time on each single attempt and required fewer attempts in total to be satisfied with their performance.

On average the trainees were satisfied with a *Best training score* of 74% before deciding to take the test scenario. Most commonly the trainees performed the *Test scenario* within the same day as their last *Training scenario* attempt. As shown in Table 5 the average *Test scenario score* was 83% of the full
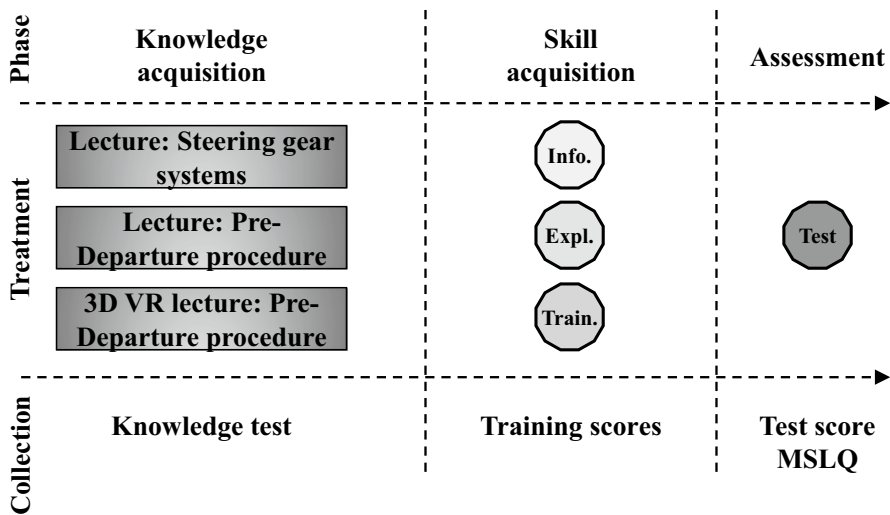

Fig. 6 Experiment procedure

Table 5 Descriptive statistics

Variable	N	Mean	SD	Shapiro-Wilk p
Knowledge test	18	8.67	1.49	0.002
Total training attempts	18	2.61	1.33	0.029
Total training time (s)	18	4550	1853	0.817
Best training scenario score	18	163.33	59.60	0.004
Test scenario score	18	183.61	39.40	0.044
Time Demand (min)	18	90.69	49.35	<0.001
Time Retention (min)	18	30.06	42.91	0.098
Progress Rate	18	3.12	1.65	0.002
Extraversion	17	5.13	0.97	0.634
Agreeableness	17	5.76	0.65	0.927
Conscientiousness	17	5.08	0.97	0.475
Emotional Stability	17	5.42	0.80	0.860
Openness to Experience	17	4.94	0.71	0.118
IGO	18	5.19	0.97	0.922
EGO	18	4.72	1.03	0.464
TV	18	5.32	1.06	0.001
CLB	18	5.67	0.76	0.100
SELP	18	4.66	0.95	0.257
TA	18	4.83	0.91	0.960
M_{POOLED}	18	5.02	0.63	0.516

IGO = Intrinsic Goal Orientation, EGO = Extrinsic Goal Orientation, TV = Task Value, CLB = Control of Learning Beliefs, SELP = Self-Efficacy for Learning and Performance, TA = Test Anxiety, M_{POOLED} = Pooled Motivation Scales

possible score. The quantifications of the collected simulator metrics show that on average 90 min was needed to train enough to obtain a full 100% score (*Time Demand*); however, they chose to stop 30 min before that would occur (*Time Retention*). Differences in the variables *Total training attempts*, *Total training time*, and performance score of each single *Training scenario* attempt could only describe the learning process when quantified as the *Progress Rate*. Displaying very different learning processes, the range of the trainee's *Progress Rates* was from 1 to 7 score points per accumulated minute engaged in training, with an average at 3.12. For example, the 4 fastest performers had a *Progress Rate* between 4 and 7, spent either one or two attempts on the *Training scenario*, and scored above 83% on their last attempt. Interestingly, the group with a *Progress Rate* between 1 and 3 contains the poorest training score performers as well as several performers above 85%.

As one of the motivation scales and most of the simulator performance indicators was found to be non-normally distributed, non-parametric tests were selected for the correlation and related sample testing.

4.1 How is motivation and personality related in this learning process?

Hypothesis 1 stated that the personality trait factors correlate with motivation scales. As *Task Value* revealed to have a non-normal distribution, the non-parametric Spearman's rho correlation test was run, as summarized in Table 6. *Extraversion* resulted to be significantly correlated with *Intrinsic Goal Orientation* and *Task Value*. As did *Agreeableness*, which in addition shows a moderate, but non-significant correlation with *Self-Efficacy for Learning and Performance*. *Conscientiousness* had moderate but non-significant correlation with *Intrinsic Goal Orientation*, *Task Value*, and *Control of Learning Beliefs*. *Openness* had a moderate but non-significant correlation with *Intrinsic Goal Orientation* and a negative correlation with *Test Anxiety*. Hypothesis 1 was accepted for the significant correlations *Extraversion — Intrinsic Goal Orientation*, *Extraversion — Task Value*, *Agreeableness — Intrinsic Goal Orientation*, and *Agreeableness — Task Value*.

Hypothesis 2 stated that personality trait factors are facilitators of motivation. As 4 correlations between personality and motivation were established by hypothesis 1, the predictive features of the BFI factors were investigated on the *Pooled Motivation Scales*. A stepwise linear regression was run and resulted in *Extraversion* being the single significant personality factor that can predict the index of the *Pooled Motivation Scales* ($F(1,16) = 4.907$, $p = 0.043$, $R^2 = 0.246$). The model was assessed with a modified Breusch-Pagan test ($\chi^2(1,17) = 0.556$, $p = 0.456$) and a White test ($\chi^2(2,17) = 1.145$, $p = 0.564$) with both unable to detect heteroskedasticity or non-linearity of error residuals. Normal distribution of the Standardized Residuals ($W = 0.962$, $p = 0.670$) was confirmed with the Shapiro-Wilk test. Hypothesis 2 was accepted within the limitation of the regression model.

Table 6 Spearman's rho correlation between BFI factors and motivation scales

	Extraversion	Agreeableness	Conscientiousness	Emotional Stability	Openness to Experience
IGO	**0.508***	**0.546***	**0.481**	0.275	**0.433**
EGO	0.260	0.251	0.139	0.204	−0.049
TV	**0.604***	**0.580***	**0.419**	0.269	0.301
CLB	−0.075	0.196	**0.433**	−0.048	0.139
SELP	0.334	**0.404**	0.186	0.040	0.293
TA	0.088	−0.143	−0.341	0.285	**−0.416**
M_{POOLED}	**0.546***	**0.447**	0.300	0.207	0.239

ρ: 0.4 < **moderate** > 0.6 < **strong** > 0.8

IGO = Intrinsic Goal Orientation, EGO = Extrinsic Goal Orientation, TV = Task Value, CLB = Control of Learning Beliefs, SELP = Self-Efficacy for Learning and Performance, TA = Test Anxiety, MPOOLED = Pooled Motivation Scales

*Correlation is significant at the 0.05 level (2-tailed)

4.2 Is test performance a product of training performance or training process?

For overview, Table 7 displays the correlations between the simulator data. The *Knowledge test* has no correlation to the performance indicators or the training quantifications. *Total training* attempts correlate with *Total training time*, which is to be expected. As the training quantifications are based on *Total training attempts*, *Total training time*, and the *Best training score* these relationships appear as correlations in the table, but have to be disregarded as they are dependent. The correlation between the *Best training score* and the *Test scenario score* indicates that there is basis for comparing these as related samples. Interestingly, there is a significant correlation between the *Test scenario score* and the *Time Retention*. As to be expected the demographic data and simulator data were not normally distributed (see Table 5).

Hypothesis 3 stated that training performance will be different than test performance, considering the absence of all supportive structures in the *Test scenario* level. As non-normally distributed, the *Best training score* was compared with the *Test scenario score* through non-parametric tests. Both the Wilcoxon Signed Rank test ($Z = -1.441$, $p = 0.149$) and the Friedman's Two-Way ANOVA ($\chi^2_F(1) = 1.923$, $p = 0.166$) failed to detect a significant difference; thus, hypothesis 3 was rejected.

Hypothesis 4 stated that the test performance will be predicted by the training process. Of all the training performance indicators and the training quantifications, a stepwise linear regression returned *Time Retention* as the single significant predictor of the *Test scenario score*, $F(1,17) = 12.400$, $p = 0.003$, $R^2 = 0.437$. The model was assessed with a modified Breusch-Pagan test ($\chi^2(1,18) = 4.199$, $p = 0.040$) and a White test ($\chi^2(2,18) = 7.076$, $p = 0.029$) which find heteroskedasticity of the residual errors and possible non-linearity. Normal distribution of the Standardized Residuals ($W = 0.958$, $p = 0.557$) was confirmed with the Shapiro-Wilk test. As the model was in breach with the assumptions of a linear regression the statistical significance was not trusted and hypothesis 4 was rejected.

4.3 Is motivation related to the training process and training performance, or test performance?

Hypothesis 5 stated that motivation scales would correlate with the simulator metrics. The correlations found in Table 8 show that the *Knowledge test* has a moderate negative correlation with *Extrinsic Goal Orientation* and a significant negative correlation with *Task Value*. *Extrinsic Goal Orientation* also has a moderate correlation with *Total training time*. The *Best training score* has a significant correlation with *Self-Efficacy for Learning and Performance* and a moderate negative correlation with *Test Anxiety*. Hypothesis 5 was accepted for the significant correlations *Task Value — Knowledge test* and *Self-Efficacy for Learning and Performance — Best training score*.

Hypothesis 6 stated that training performance will be predicted by motivation. A stepwise linear regression was run and resulted in *Self-Efficacy for Learning and*

Table 7 Spearman's rho correlation between training performance indicators, training quantifications, and test performance

	Knowledge test	Total training attempts	Total training time	Best training score	Time Demand	Time Retention	Progress Rate	Test score
Knowledge test	1							
Total training attempts	−0.165	1						
Total training time	−0.260	**0.808****	1					
Best training score	−0.039	−0.006	0.066	1				
Time Demand	0.037	**0.452**	0.290	**−0.636****	1			
Time Retention	0.014	0.135	0.015	**−0.938****	**0.803****	1		
Progress Rate	−0.048	**−0.440**	−0.285	**0.639****	**−0.999****	**−0.808****	1	
Test score	0.265	0.117	0.192	**0.538***	−0.087	**−0.474***	0.088	1

ρ: 0.4 < **moderate** > 0.6 < **strong** > 0.8

**Correlation is significant at the 0.01 level (2-tailed)

*Correlation is significant at the 0.05 level (2-tailed)

Table 8 Spearman's rho correlation between motivation scales and training performance indicators, training quantification, and test performance

	IGO	EGO	TV	CLB	SELP	TA
Knowledge test	−0.352	**−0.463**	**−0.605****	−0.138	−0.336	−0.020
Total training attempts	0.012	0.224	0.019	0.013	0.353	0.156
Total training time	0.050	**0.404**	0.209	0.112	0.346	0.227
Best training score	0.263	0.362	0.269	−0.072	**0.471***	**−0.420**
Time Demand	−0.198	−0.168	−0.160	0.136	−0.044	0.080
Time Retention	−0.205	−0.256	−0.204	0.102	−0.331	0.383
Progress Rate	0.201	0.164	0.166	−0.136	0.054	−0.084
Test scenario score	−0.064	−0.049	0.000	0.163	0.220	−0.399

ρ: 0.4 < **moderate** > 0.6 < **strong** > 0.8

IGO = Intrinsic Goal Orientation, EGO = Extrinsic Goal Orientation, TV = Task Value, CLB = Control of Learning Beliefs, SELP = Self-Efficacy for Learning and Performance, TA = Test Anxiety

**Correlation is significant at the 0.01 level (2-tailed)

*Correlation is significant at the 0.05 level (2-tailed)

Performance being the only significant motivational factor that can predict the *Best training score* ($F(1,17) = 5.323$, $p = 0.035$, $R^2 = 0.250$). The model was assessed with a modified Breusch-Pagan test ($\chi^2(1,18) = 0.851$, $p = 0.356$) and a White test ($\chi^2(2,18) = 5.467$, $p = 065$) both unable to detect heteroskedasticity or non-linearity of error residuals. Normal distribution of the Standardized Residuals ($W = 0.928$, $p = 0.182$) was confirmed with the Shapiro-Wilk test. Hypothesis 6 was accepted within the limitation of the regression model.

5 Discussion

The statistical analysis resulted in the hypotheses for research questions 1 and 3 to be accepted whereas hypotheses for research question 2 were rejected, as summarized in Table 9.

Hypothesis 1 finds some personality trait factors that correlate with motivation scales. The findings could suggest that a personality profile with high *Extraversion* will be beneficial for recognizing value in the challenge of solving the task itself (correlating with *Intrinsic Goal Orientation*), as well as the value of the content on the task (correlating with *Task Value*). This is complementary to present scientific knowledge (De Feyter et al., 2012; Shiota et al., 2006). *Agreeableness* also addresses the same motivational beliefs, suggesting the personality profile to have some elements of subservient disposition or succumbing adaption of accepting the task as a learning outcome. *Conscientiousness* offers some correlations and is the only trait correlated to *Control of learning Beliefs*. Interestingly, *Emotional Stability*, sometimes referred to as neuroticism, shows no meaningful correlation to motivation, suggesting that the trait is either neutral or irrelevant for motivation. *Openness to Experience* shows only moderate correlations, the interesting one being a negative

Table 9 Research questions and hypothesis inference

	Statement	Inference
RQ1	How is motivation and personality related in this learning process?	
H1	*Personality trait factors correlate with motivation scales.*	Accepted
H2	*Personality trait factors are facilitators of motivation.*	Accepted
RQ2	Is test performance a product of training performance or training process?	
H3	*Training performance will be different than Test performance.*	Rejected
H4	*Test performance will be predicted by the training process.*	Rejected
RQ3	Is motivation related to the training process and training performance, or test performance?	
H5	*Motivation scales correlate with the simulator metrics.*	Accepted
H6	*Training performance will be predicted by motivation.*	Accepted

correlation to *Test Anxiety*. This can suggest that openness is a trait that provides resilience against the cognitive element of anxiety or has a mitigating effect on the emotional element of anxiety.

Viewing motivation scaled as an accumulative composite construct, *Extraversion* is suggested to be the strongest predictor of the *Pooled Motivation Scales*. From hypothesis 2 a regression model was constructed where the level of *Extraversion* explains 24.6% of the difference in trainees' motivation. Regardless of the small sample size the effect found in the model is robust enough to be considered.

In total the findings of research question 1 can be interpreted to indicate the existence of a personality profile that is favourable for motivation with novel students in a maritime education context, which is strongest with high *Extraversion*.

As to be expected there was a correlation between the *Best training score* and the *Test scenario score*. Hypothesis 3 predicted that these scores would be at a different performance level due to the difference in conditions, which proved to be insignificant. The descriptive statistics reveal that the score increases between these conditions, even as the programmed supportive structures of the *Training scenario* is completely absent in the *Test scenario*. Further it is noted that the central tendency consolidates as evident in the coefficient of variance for the *Best training score* and the *Test scenario score*, which was 36.5% and 21.4%, respectively. Although not captured by the statistical testing, this indicates that there is something happening with the learning process between activities, and that it is converging across the sample (Kraiger et al., 1993).

Assuming the performance of a repeated training process to be approximately linear (Kelley, 1969), quantifications of the total required temporal effort (*Time Demand*), the remaining temporal effort (*Time Retention*), and the linear *Progress Rate* were made from the simulator metrics collected. Independent of the *Test scenario* metrics (*Test scenario score*), these quantifications along with the other simulator variables were investigated as causal to the final *Test scenario score*. The regression model of hypothesis 4 returned *Time Retention*, i.e. the remaining time to achieve a full *Training scenario* score (*Best training score*), as the single significant predictor, explaining 43.7% of the variance in *Test scenario scores*. If

heteroskedasticity is present, the coefficient estimates of the regression are still valid; however, the model output and *p*-value might be wrong as calculated from an incorrect standard error. The regression models in hypotheses 2 and 6 do not raise this concern. However, the regression of hypothesis 4 needed to be evaluated and as it fails to meet the assumptions of a linear regression the effect and significance are unvalidated. It is likely to believe that with a small sample study such data distortion can occur, especially when one end of the simulator score range can be viewed as a latent performance goal. Although no such goal was given by instruction, there was no prevention of the trainees setting a full score as their individual goal.

In total the statistics collected for research question 2 cannot determine with certainty that the *Test scenario* performance (*Test scenario score*) is a product of the point in time where training ceases (*Time Retention*), the path to that point (*Progress Rate*), or both. The investigation shows that there is something happening to the trainees learning process, between the last *Training scenario* attempt and the *Test scenario*, which was not clearly captured by the measurements in this study. Interestingly, the *Knowledge test* at the end of the knowledge acquisition phase shows no relation to any measurement of the skill acquisition phase. This could suggest that the effect of prior knowledge on skill-based performance is attenuated through training repetition (Kraiger et al., 1993).

All correlations between the *Knowledge test* and the motivation scales were negative, suggesting that the knowledge level prior to the simulator training is irrelevant for motivational beliefs of the total learning process. The moderate correlation between *Total training time* and *Extrinsic Goal Orientation* might reveal that perceiving value in one's training performance can affect the effort provided to obtain one's goal.

The *Best training score* correlated with *Self-Efficacy for Learning and Performance*, and *Test Anxiety*, which indicates a calibration of the trainee's self-perception and confidence through the repeated training. The regression model of hypothesis 6 shows that 25% of the training score difference was predicted by self-efficacy. De Feyter et al. (2012) find that high self-efficacy supports academic performance and that it is developed through the learning process.

In sum, research question 3 finds a relation between motivation and training performance. This could support that in a repeated training design motivation through self-efficacy relates to the training performance as a product of calibrating the confidence in one's ability to perform (Zimmerman, 2008).

5.1 Post hoc follow-up

As hypothesis 3 finds the *Best training score* and the *Test scenario score* significantly correlated, but not with significant difference of the ranked means, the question of this relationship being a design flaw arose. Testing hypothesis 4 indicated that there was some prediction of the test performance (*Test scenario score*) based on measures of the training activity (*Time Retention*). If the correlation was an indication of causation, a design flaw could be the cause of the *Test scenario scores* if no other model than such direct effect was found ($F(1, 16) = 7,406$, $R^2 = 0,316$, $p =$

0.015). As hypothesis 6 found relationships between the *Best training score* and one motivation scale (*SELP*), motivation was suspected to explain some of this relationship between the training and test performances.

A floodlight moderator analysis was performed regressing the *Best training scores* on the *Test scenario scores* with the *Pooled Motivation Scales* as a moderator (Model 1; Hayes, 2022), which resulted in a significant overall model, $F(3, 14) = 9.254$, $R^2 = 0.665$, $p < 0.001$ (see Fig. 7). A significant interaction $F(1, 14) = 9.880$, $R^2_{change} = 0.237$, $p = 0.007$ explained that 23.7% of the model's effect was due to the *Pooled Motivation Scales* moderating the relationship. Conditional effects of the *Best training scores* at the different levels of the moderator (*Mn* = 5.037, *SD* = 0.616) reveal a Johnson-Neyman point at a *Pooled Motivation Scales* level of 4.632 and above ($B_{JN} = 0.272$, *SE* = 0.127, $t(1, 14) = 2.245$, $p = 0.05$). This means that the trainees (77.7%) who experienced motivation above the JN point had a significant positive effect of motivation on performance, when the conditions were, by their own choice, changed from a *Training scenario* to a *Test scenario* condition.

The post hoc results help to explain the study with an abstract point of view based on hypothesized relationships explored that indicated mechanisms of the constructs under investigation. Figure 7 helps connect research questions 2 and 3 where 66.5% of *Test scenario* performance is a product of the interaction between *Training scenario* performance and experienced motivational stimuli.

5.2 Implications

The prediction of motivation based on *Extraversion* was made using a short version (Engvik & Clausen, 2011) of the BFI-44 (John & Srivastava, 1999). This provide support for the possibility of using shorted BFI scales at an indicative level if the alpha reliabilities are in order.

This study shows that CBS could be used as a tool in an integrated learning process with training of procedural skills. It can deliver repeated training at the trainee's discretion and perhaps be used as part of educational courses or training programs along with different learning activities. In view of incorporating this CBS platform
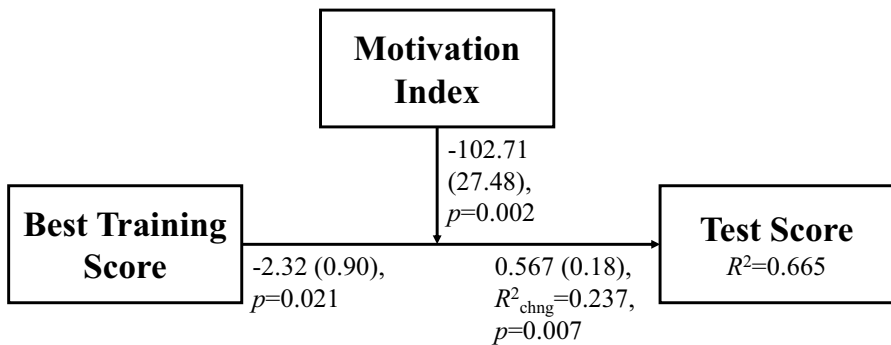

Fig. 7 Moderator model with unstandardized regression weights, standard errors, effect sizes, and path significance

as part of a smart learning environment, it holds some technical desirable characteristics such as engaging, flexible, adaptive, and personalized (Spector, 2014). Although it can be a convenient platform for the trainee, the CBS platform would require different tasks and support from the simulator instructor than with the traditional simulators.

One common observation of effectiveness was that time engaged decreased to an individual saturation point as the trainees improved their performance. Some trainees exhibited an inflection point in their training progression where the first part of their training yielded slow progress until the slope started to increase. Training progress pattern has shown to be linear with adaptive systems and follows an *s*-shaped curve when fixed (Kelley, 1969), and training performance across a sample is expected to converge over time (Kraiger et al., 1993).

Contributions to the simulator instructor here are the indication of importance in facilitating motivation through designing the training delivery and potential use of the training quantifications made. Building a foundation of motivation through repetitive training with performance feedback clearly indicates a positive prediction of exam performance, given the circumstances of this study. The quantifications describe the trainee's learning process beyond simple performance indicators and pinpoint where in the process they currently stand. This is relevant for *feedforward* instruction and planning the following training need. With future dedicated research, specific economic points for training conclusion could be identified that better predict training transfer results.

To facilitate engagement with the learning program, design features to optimize individual mental resources were incorporated such as personalization of the training, virtual agents representing instructional functions, automated feedback, and vocational spacing of phases and content (Clark, 2021). Trainees might have to adopt new learning patterns that integrate formal and informal learning through a more encompassing process across location, time availability, and interaction with people and technology (Tabuenca et al., 2021). The social interactions of the education could be considered in the CBS training design to both involve one-on-one instruction (Bloom, 1984) and the benefits of peer instruction (Crouch & Mazur, 2001), and peer comparison (Vesin et al., 2018).

Clark (2021) argues three core instructional principles for virtual environment learning to be effective: (1) minimize extraneous mental load, (2) optimize germane mental load, and (3) provide feedback. Adaptive and personalized maritime CBS training can in design address these principles of optimizing the presented content for engagement and provide effective feedback. Asynchronous training as done in this study could reduce some external mental load, compared to traditional synchronous training. As participation is individual and in the surroundings of the trainee's preference, there is low or no social presence with instructors or peers, and the trainee controls the pace. Positive learning effects could be obtainable with virtual agents providing automated feedback, automated assessment, and automated actions that change the simulation scenario. The developers of CBS technology could further develop current models to facilitate features of adaptive training design with real-time performance indication, indication of progress, and feedback possibilities that substitute the supervision of a present or remote instructor. The challenge of

authoring scenarios can be met by team effort or single dedicated instructors; however, servicing all necessary capacities for programming requires some resources and dedication. Integrating CBS training with maritime educational programs to target certain learning outcomes at the knowledge and skill level is technologically and pedagogically feasible; however, although CBS can offer supplemental contribution to such educations it cannot fully substitute all features of the traditional on-campus simulators.

5.3 Limitations

A larger sample could have prevented the data distortion experienced in the study. Small sample studies could raise some concerns for the statistical robustness of the hypothesis tests used and the resulting effect sizes.

A test's statistical significance is not less valid with a small sample size; the implication is that the observed effect size needs to be larger because it will be harder to distinguish between a real effect and random variation (Hackshaw, 2008). The concern for statistical significance of small sample studies, prior to conducting the study itself, is that the confidence intervals will be larger and that the test probability becomes more fragile (Rosenthal, 1991). This because the significance test is a product of the size of the effect and the size of the study.

Another concern of small sample studies often refers to the statistical power achieved by (1) the observed effect size based on the sample size, (2) the defined α probability level, and (3) the sample size (Field & Hole, 2003). In practice, the recommended sample size for a test is found a priori by a hypothetical population effect size, the defined α probability level, and the statistical power one wants the test to have (which is the probability of finding an effect if there is one). The observed statistical power post hoc a test will be the probability of detecting an effect if there truly is one, when replicating the study with the same sample size and expecting the same effect size.

The effect size of the regression models of hypotheses 2 and 6 categorizes as large (Cohen, 1992), which should be favourable for model fitness as the sample is considered small in absolute terms and as the effect size is based on the small sample size (Field, 2009). The accuracy of correlations could also suffer with small samples (Schönbrodt & Perugini, 2013). Even as parametric tests could be applied (Norman, 2010) the non-parametric tests used are an acceptable approach with the current sample size and the assumptions made. As non-parametric ranking tests are less powerful they are more likely to not detect an effect when there is one (Field & Hole, 2003), rendering the choice of tests the most conservative alternative. Hackshaw (2008) argues that smaller samples are sufficient in hypothesis-generating studies if statistical power is supported. Regarding generalizability, empirical studies should have a clearly defined target population and the sample should be described to such an extent that it is possible to replicate (Simons et al., 2017). As such, we cannot with this study alone claim external validity of our interpretations to populations beyond the sample of our target population (Stroebe et al., 2018).

Although the full MSLQ was collected, the Self-Regulated Learning (SRL) section of the questionnaire was not included in the study and analysis. Self-regulation and metacognitive strategies hold low measurement validity with novice students when not integrated with or preceded by a training program that target such outcomes. This was evident by low internal consistency of the SRL scales and inclusion of these to the study would be a leap beyond statistical soundness.

Reliability indexes for the measured variables *Agreeableness*, *Emotional Stability*, *Openness*, *EGO*, and *TA* were found lower than what is considered an acceptable internal consistency and lower than expected from the generalization studies in Table 2. This means that interitem correlation of the variable scales was low due to variance between item responses, or due to less responses than necessary to establish such correlation (DeVellis, 2017). The implication of this is the risk of erroneous interpretation or replicability of statistical results. However, the risk was operationally mitigated to the possibility of not "providing the full picture" as these variables were not included in inference of results or interpreted generalizable to any population.

6 Conclusion and future research

This paper employed a quasi-experimental study to investigate trainee motivation, personality traits, and task performance using a novel training design. The study focused on the trainees and was designed integrated with the relevant education courses through the phases (1) lecture-based knowledge acquisition, (2) asynchronous skill acquisition, and (3) performance assessment. This research found the personality factor Extraversion to explain some occurrence of motivation, and one facet of motivation to predict performance in training. Further contemplation depicted a model where the relationship between performance in a training condition and a test condition was positively moderated by motivation.

With cloud-based simulation training, we might have helped to mitigate trainee characteristics by facilitating positive motivational factors in their learning process. The findings indicated certain personality traits as favourable for motivation and facets of motivation that are positively related to the training process. For test condition performance, motivation is a factor that should be considered to positively interact with previous training. Leveraging this could help simulator instructors designing and delivering training.

Training progress patterns should be further investigated with the application of CBS to design training delivery that address and mitigate differences in trainee characteristics.

It would be a prominent follow-up to repeat a similar experiment with a control group for both the knowledge- and skill-acquisition phases and expand the study with a knowledge test after the simulator training was complete for transfer of knowledge measurement. Also administering the MSLQ before and during the training phase as the development of motivational beliefs might hold insight into how CBS training delivery and design can be improved.

With the present development of increasingly immersive simulators and enhanced technical flexibility, adaptive and personalized training systems could be a prominent approach for simulator developers to consider that might improve learning and trainee performance. At the current stage, the position of CBS in maritime education and training is believed to best serve as a supplement to the overall traditional learning process. CBS is ready for such integration, which could enhance outcomes by expanding the applicability of simulator training, but not yet substitute the traditional stationary technologies. The authors welcome further research to contribute with new knowledge in this relatively unpaved context of maritime training research.

Declarations

Conflict of interest The authors declare no competing interests.

References

Barrick MR, Mount MK, Judge TA (2001) Personality and performance at the beginning of the new millennium: what do we know and where do we go next? *Int J Sel Assess 9*(1&2):9–30. https://doi.org/10.1111/1468-2389.00160

Bell B, Tannenbaum SI, Ford JK, Noe RA, Kraiger K (2017) 100 years of training and development research: what we know and where we should go. J Appl Psychol 102:305. https://doi.org/10.1037/apl0000142

Bloom BS (1984) The 2 sigma problem: the search for methods of group instruction as effective as one-to-one tutoring. *Educ Res 13*(6):4–16. https://doi.org/10.3102/0013189x013006004

Chernikova O, Heitzmann N, Stadler M, Holzberger D, Seidel T, Fischer F (2020) Simulation-based learning in higher education: a meta-analysis. Rev Educ Res 90(4):499–541. https://doi.org/10.3102/0034654320933544

Clark RC (2021) Multimedia learning in e-courses. In: Fiorella L, Mayer RE (eds) *The Cambridge Handbook of Multimedia Learning*, 3rd edn. Cambridge University Press, pp 537–551. https://doi.org/10.1017/9781108894333.054

Cohen J (1992) A power primer. Psychol Bull 112(1):155–159. https://doi.org/10.1037//0033-2909.112.1.155

Colquitt JA, LePine JA, Noe RA (2000) Toward an integrative theory of training motivation: a meta-analytic path analysis of 20 years of research. J Appl Psychol 85(5):678–707. https://doi.org/10.1037/0021-9010.85.5.678

Crouch CH, Mazur E (2001) Peer instruction: ten years of experience and results. Am J Phys 69(9):970–977. https://doi.org/10.1119/1.1374249

Database for Statistics on Higher Education (2020) https://dbh.nsd.uib.no/. Accessed 12 Sept 2021

De Feyter T, Caers R, Vigna C, Berings D (2012) Unraveling the impact of the Big Five personality traits on academic performance: the moderating and mediating effects of self-efficacy and academic motivation. Learn Individ Differ 22(4):439–448. https://doi.org/10.1016/j.lindif.2012.03.013

DeVellis RF (2017) Scale development: theory and applications, 4th edn. SAGE

DNV (2021) Maritime simulator systems. In DNV (ed) DNV-ST-0033, vol. DNV-ST-0033, p, 171

DNV (2022) Competence of teaching professionals. In DNV (ed) DNV-ST-0024

Duncan T, Pintrich P, Smith D, McKeachie W (2015) Motivated Strategies for Learning Questionnaire (MSLQ) manual. National Center Res Improve Postsecondary Teach Learn 2:2547–6968. https://doi.org/10.13140/RG.2.1.2547.6968

Dunning D (2018) The best option illusion in self and social assessment. Self and Identity 18:1–14. https://doi.org/10.1080/15298868.2018.1465460

Engvik H, Clausen SE (2011) Norsk kortversjon av Big Five Inventory (BFI-20). Tidsskr Nor Psyk 48:869–872. https://psykologtidsskriftet.no/oppsummert/2011/09/norsk-kortversjon-av-big-five-inventory-bfi-20. Accessed 20 May 2021

Feldon DF, Jeong S, Clark RE (2021) Fifteen common but questionable principles of multimedia learning. In: Fiorella L, Mayer RE (eds) The Cambridge Handbook of Multimedia Learning, 3rd edn. Cambridge University Press, pp 25–40. https://doi.org/10.1017/9781108894333.005

Field A (2009) Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll), 3rd edn. SAGE

Field A, Hole G (2003) How to design and report experiments. SAGE

Grossman R, Heyne K, Salas E (2014) Game- and simulation-based approaches to training. In: The Wiley Blackwell Handbook of the Psychology of Training, Development, and Performance Improvement, pp 205–223. https://doi.org/10.1002/9781118736982.ch12

Gully S, Chen G (2010) Individual differences, attribute–treatment interactions, and training outcomes. In: Kozlowski SWJ, Salas E (eds) Learning, Training, and Development in Organizations. Routledge / Taylor & Francis Group, pp 3–64

Gurtner J-L (2014) Effective virtual learning environments. In: The Wiley Blackwell Handbook of the Psychology of Training, Development, and Performance Improvement, pp 188–204. https://doi.org/10.1002/9781118736982.ch11

Hackshaw A (2008) Small studies: strengths and limitations. Eur Respir J 32(5):1141–1143. https://doi.org/10.1183/09031936.00136408

Hattie J, Timperley H (2016) The power of feedback. Rev Educ Res 77(1):81–112. https://doi.org/10.3102/003465430298487

Hayes AF (2022) Introduction to mediation, moderation, and conditional process analysis : a regression-based approach, 3rd edn, Guilford Publications

Holland DF, Kraha A, Zientek LR, Nimon K, Fulmore JA, Johnson UY, Ponce HF, Aguilar MG, Henson RK (2018) Reliability generalization of the Motivated Strategies for Learning Questionnaire: a meta-analytic view of reliability estimates. SAGE Open 8(3):215824401880233. https://doi.org/10.1177/2158244018802334

John OP, Srivastava S (1999) The Big Five Trait taxonomy: history, measurement, and theoretical perspectives. In: Handbook of personality: Theory and research, 2nd edn. Guilford Press, pp 102–138

Johnson CI, Marraffino MD (2021) The feedback principle in multimedia learning. In: Fiorella L, Mayer RE (eds) The Cambridge Handbook of Multimedia Learning, 3rd edn. Cambridge University Press, pp 403–417. https://doi.org/10.1017/9781108894333.042

Kanfer R, Ackerman PL (1989) Motivation and cognitive abilities: an integrative/aptitude-treatment interaction approach to skill acquisition. J Appl Psychol 74(4):657–690. https://doi.org/10.1037/0021-9010.74.4.657

Kelley CR (1969) What is adaptive training? Human Factors 11(6):547–556. https://doi.org/10.1177/001872086901100602

Kim T-E, Sharma A, Bustgaard M, Gyldensten WC, Nymoen OK, Tusher HM, Nazir S (2021) The continuum of simulator-based maritime training and education. WMU J Marit Aff 20(2):135–150. https://doi.org/10.1007/s13437-021-00242-2

Kinshuk. (2016) Designing adaptive and personalized learning environments. Routledge. https://doi.org/10.4324/9781315795492

Kraiger K, Ford JK, Salas E (1993) Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. J Appl Psychol 78(2):311–328. https://doi.org/10.1037/0021-9010.78.2.311

Kulhavy RW (1977) Feedback in written instruction. Rev Educ Res 47(2). https://doi.org/10.3102/00346543047002211

Kvinge LMR, Engelsen KS (2016) Omsetjing og tilpassing av MSLQ: IMTE-rapport nr. 1/16

Lajoie SP (2021) Multimedia learning with simulations. In: Fiorella L, Mayer RE (eds) *The Cambridge Handbook of Multimedia Learning*, 3rd edn. Cambridge University Press, pp 461–471. https://doi.org/10.1017/9781108894333.048

Landsberg CR, Astwood RS, Van Buskirk WL, Townsend LN, Steinhauser NB, Mercado AD (2012) Review of adaptive training system techniques. Mil Psychol 24(2):96–113. https://doi.org/10.1080/08995605.2012.672903

Lindner MA (2021) Principles for educational assessment with multimedia. In: Fiorella L, Mayer RE (eds) *The Cambridge Handbook of Multimedia Learning*, 3rd edn. Cambridge University Press, pp 552–565. https://doi.org/10.1017/9781108894333.055

Makransky G (2021) The immersion principle in multimedia learning. In: Fiorella L, Mayer RE (eds) *The Cambridge Handbook of Multimedia Learning*, 3rd edn. Cambridge University Press, pp 296–303. https://doi.org/10.1017/9781108894333.031

Mallam SC, Nazir S, Renganayagalu SK (2019) Rethinking maritime education, training, and operations in the digital era: applications for emerging immersive technologies. J Mar Sci Eng 7(12). https://doi.org/10.3390/jmse7120428

Mayer RE (2021) Cognitive theory of multimedia learning. In: Fiorella L, Mayer RE (eds) *The Cambridge Handbook of Multimedia Learning*, 3rd edn. Cambridge University Press, pp 57–72. https://doi.org/10.1017/9781108894333.008

Motejlek J, Alpay E (2021) Taxonomy of virtual and augmented reality applications in education. IEEE Trans Learn Technol 14(3):415–429. https://doi.org/10.1109/tlt.2021.3092964

Norman G (2010) Likert scales, levels of measurement and the "laws" of statistics. Adv Health Sci Educ 15(5):625–632. https://doi.org/10.1007/s10459-010-9222-y

O'Donnell E, Lawless S, Sharp M, Wade VP (2015) A review of personalised e-learning. Intern J Dist Educ Technol 13(1):22–47. https://doi.org/10.4018/ijdet.2015010102

Pajares F (2012) Motivational role of self-efficacy beliefs in self-regulated learning. In: *Motivation and self-regulated learning: Theory, research, and applications*. Routledge, pp 111–139

Parks L, Guay RP (2009) Personality, values, and motivation. Personal Individ Differ 47(7):675–684. https://doi.org/10.1016/j.paid.2009.06.002

Parong J (2021) Multimedia learning in virtual and mixed reality. In: Fiorella L, Mayer RE (eds) *The Cambridge Handbook of Multimedia Learning*, 3rd edn. Cambridge University Press, pp 498–509. https://doi.org/10.1017/9781108894333.051

Pintrich PR, Smith DAF, Garcia T, McKeachie W (1991) A manual for the use of the motivated strategies for learning questionnaire (MSLQ)

Radianti J, Majchrzak TA, Fromm J, Wohlgenannt I (2020) A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda. Comput Educ 147:85. https://doi.org/10.1016/j.compedu.2019.103778

Rosenthal R (1991) *Meta-analytic procedures for social research*. Sage Publications. https://doi.org/10.4135/9781412984997

Salas E, Tannenbaum SI, Kraiger K, Smith-Jentsch KA (2012) The science of training and development in organizations. Psychol Sci Public Interest 13(2):74–101. https://doi.org/10.1177/1529100612436661

Schönbrodt FD, Perugini M (2013) At what sample size do correlations stabilize? J Res Pers 47(5):609–612. https://doi.org/10.1016/j.jrp.2013.05.009

Schrader C, Kalyuga S, Plass JL (2021) Motivation and affect in multimedia learning. In: Fiorella L, Mayer RE (eds) *The Cambridge Handbook of Multimedia Learning*, 3rd edn. Cambridge University Press, pp 121–131. https://doi.org/10.1017/9781108894333.012

Sharma A, Nazir S (2021) Assessing the technology self-efficacy of maritime instructors: an explorative study. Educ Sci 11(7). https://doi.org/10.3390/educsci11070342

Shiota MN, Keltner D, John OP (2006) Positive emotion dispositions differentially associated with Big Five personality and attachment style. J Posit Psychol 1(2):61–71. https://doi.org/10.1080/17439760500510833

Simons DJ, Shoda Y, Lindsay DS (2017) Constraints on Generality (COG): a proposed addition to all empirical papers. Perspect Psychol Sci 12(6):1123–1128. https://doi.org/10.1177/1745691617 708630

Spector JM (2014) Conceptualizing the emerging field of smart learning environments. Smart Learn Environ 1(1):1. https://doi.org/10.1186/s40561-014-0002-7

Stroebe W, Gadenne V, Nijstad BA (2018) Do our psychological laws apply only to college students?: external validity revisited. Basic Appl Soc Psychol 40(6):384–395. https://doi.org/10.1080/01973 533.2018.1513362

Suh A, Prophet J (2018) The state of immersive technology research: a literature analysis. Comput Hum Behav 86:77–90. https://doi.org/10.1016/j.chb.2018.04.019

Tabuenca B, Serrano-Iglesias S, Carruana-Martin A, Villa-Torrano C, Dimitriadis YA, Asensio-Perez JI, Alario-Hoyos C, Gomez-Sanchez E, Bote-Lorenzo ML, Martinez-Mones A, Delgado Kloos C (2021) Affordances and core functions of smart learning environments: a systematic literature review. IEEE Trans Learn Technol 14:1–1. https://doi.org/10.1109/tlt.2021.3067946

The Norwegian Universities and Colleges Admission Service (2020) https://www.samordnaopptak.no/ info/english/. Accessed 12 Sept 2021

Towler A, Mitchell T (2014) Facilitation in e-learning. In: *The Wiley Blackwell Handbook of the Psychology of Training, Development, and Performance Improvement*, pp 171–187. https://doi.org/10.1002/ 9781118736982.ch10

Vesin B, Mangaroska K, Giannakos M (2018) Learning in smart environments: user-centered design and analytics of an adaptive learning system. *Smart Learn Environ* 5(1):1. https://doi.org/10.1186/ s40561-018-0071-0

Zimmerman BJ (2008) Investigating self-regulation and motivation: historical background, methodological developments, and future prospects [article]. Am Educ Res J 45(1):166–183. https://doi.org/10. 3102/0002831207312909

Zimmerman BJ, Schunk DH (2012) An essensial dimension of self-regulated learning. In: Schunk BJZDH (ed) *Motivation and Self-Regulated Learning: Theory, Research and Applications*. Routledge, pp 1–30