ARTICLE

Reviewing simulator-based training and assessment in maritime education: a topic modelling approach for tracing conceptual developments

A. Camilla Wiig[1] · **Charlott Sellberg**[2] · **Mads Solberg**[3]

Abstract

This study reviews literature of simulation-based training and assessment to trace conceptual developments in maritime education and training (MET). Our aim is to highlight and examine the historical trajectory of salient topics in MET with the goal of help developing the field forward. The investigation is based on articles published from 1961 to 2021 ($n = 87$) on simulation-based training and assessment in maritime academic journals. The first review of its kind in MET, the study combines quantitative topic modelling and qualitative content analyses of exemplary texts. Our study investigates: (a) how training and assessment in simulations are conceptualised in MET research and (b) how these conceptualisations have developed over the decades. Our review reveals an educational field dominated by human factors research and dedicated to upholding industry standards. Moreover, adopting quantitative text mining we identified four prominent topics in the literature. These revolve around management of resources, academic performance, simulator training and authentic assessment. A more in-depth qualitative content analysis of the model, tracing conceptual developments over time, suggests that the word 'learning' occurs for the first time in 1980. As a concept for scientific inquiry, learning does not

✉ A. Camilla Wiig
  Camilla.wiig@usn.no

  Charlott Sellberg
  Charlott.sellberg@ait.gu.se

  Mads Solberg
  Mads.solberg@ntnu.no

1   Department of Educational Science, Faculty of Humanities, Sport, and Educational Science,
    University of South-Eastern Norway, Raveien 215, 3184 Borre, Norway

2   Department of Applied IT, Division of Learning, Communication & IT, University
    of Gothenburg, Forskningsgången 6, 41756 Göteborg, Sweden

3   Department of Health Sciences, Ålesund, Faculty of Medicine and Health Sciences,
    Norwegian University of Science and Technology, Larsgårdsveien 2, Ankeret, 6009 Ålesund,
    Norway

gain prominence in the literature on simulation in MET before 2021. Our results suggest some future directions of research and offer a methodological framework for future reviews in maritime affairs that combines quantitative topic modelling with qualitative content analysis.

1 Background and research questions

Maritime education and training (MET) exemplify a research field concerned with professional education that has undergone significant changes in recent decades. Traditionally, MET has been based on a system of apprenticeships onboard ships, whereby the practitioner became a master mariner through years of work, gradually increasing in their rank and responsibilities. Today, students become master mariners, professionals in a world of highly digitised maritime systems, through 5-year master's programmes that are regulated by international conventions and certifications (Manuel 2017). In this context, simulator-based training is not only used to bridge theory and practice, hence supplementing time spent on board to learn the profession, but also for licensing seafarers on a suite of work-related skills, for example the proficient use of radar equipment. Thus, a debate within the maritime field of research education has developed over the past 60 years. Education research has been consequential in maritime decision making while designing master's programmes through taking up simulator-based training and assessment practices as central for impactful maritime education and certification.

To better understand the emergence of simulation-based education in the maritime profession we used an unconventional review method that borrows search strategies from systematic literature review methods together with the scoping review's interest in analysing conceptual developments through qualitative analyses of academic articles (see e.g., Munn et al. 2018). In this study, we combine statistical topic modelling and qualitative analysis of central themes to examine conceptual discussions in the field's premier scientific journals. While traditional systematic literature reviews are mostly performed by doing detailed content analysis manually (Audrin and Audrin 2022), we propose using (semi) automatic methods combined with detailed content analysis. Considering the amount of relevant literature in simulator-based training and assessment, and the plurality of terms associated with this academic domain, adopting an automatic text mining technique known as topic extraction, may afford new perspectives on the research field. Using WordStat's topic extraction feature, which combines natural language processing with statistical analysis to reveal semantic structures in texts, we want to reveal the topics that define the literature i.e., the thematic characteristics of a body of text from selected scientific articles in maritime professions. Combining

automated text mining with manual qualitative content analysis we investigate in-depth the extracted 'text' for each selected key topic. By analysing the themes and their examples in the text-excerpts, matching keywords and year of publication for each case, we aim to trace the conceptual development of salient discussions developed over time within the MET literature on simulator-based training and assessment. To the best of our knowledge, no other studies provide this kind of analysis of MET. However, text mining has been used productively in reviews across a variety of professions (Thomas et al. 2011), in the history of medicine (Thompson et al. 2016) and educational research (Ferreira-Mello et al. 2019), including digital literacy (Audrin and Audrin 2022).

As far as we know, our study of simulator-based training and assessment is the first study to conduct a literature review on scientific discourse in the field of maritime education and training. By extracting articles about simulations from three domain-specific, peer-reviewed journals in MET research, the present study investigates and interprets the trajectory of conceptual developments in the academic literature about simulation-based training and assessment during the past few decades, along with how these topics have changed over time. The following research questions guide our analysis:

- RQ1: What are the most prominent topics related to training and assessment in simulation in the corpus of articles from selected maritime journals from 1961 to today?
- RQ2: What kinds of central themes are used to describe simulator-based training and assessment in the corpus, and how do these themes and their examples develop over time?

2 Methods

### 2.1 Data selection and collection

The data consist of a corpus of articles from scientific journals focusing on topics related to simulations in maritime training and education. Scientific articles can be seen as historical documents that display valuable information about the conceptual foundations that characterise a given field of research, and they offer a window into how paradigmatic assumptions in the field have developed. For that reason, the chosen journals for this review, are domain-specific and historically recognised as core publication venues within the academic study of maritime training and assessment. We chose to highlight 'training' and 'assessment' as the search words rather than explicitly searching for the concept of 'learning'. Our rationale is partly based on insights from previous reviews, and partly by doing hand search in search engines of academically known journals (Wiig 2023). In their reviews, Sellberg (2017) and Wahl and Kongsvik (2018) are documenting that the sociology of knowledge in the field of MET emphasises training and assessment, and that the research is dominated by studies of human factor interactions. This knowledge made us conduct manual hand search in the search engines of well-known journals in the field

to make a rapid exploration of how 'training' and/or 'assessment' and/or 'learning' was utilized within this domain. The terms 'training' and 'assessment' proved to be prominent concepts in headings, key words and in abstracts over three decades, while 'learning' was seldom treated as a core concept in the texts, and the term was rarely visible in headings, abstracts or key words. Since the very principle of text mining is to explore the thematic patterns specific to a corpus based on 'the frequency with which a content word appears' (Lavissière et al. 2020 p. 136), we chose to narrow the systematic search terms to 'training' and 'assessment'. These selected search words are broad enough to grasp the field of both simulator as a technology and simulation as a practice. Consequently, the corpus for our systematic analysis consists of 87 articles on simulator-based training and assessment sampled from three English, peer-reviewed and domain-specific journals within the maritime field: *Journal of Navigation*, *WMU Journal of Maritime Affairs* and *TransNav*, *The International Journal on Marine Navigation* and *Safety of Sea Transportation*. The period we have examined is from 1961, when the first article is identified in our journal search to April 2021, when the search was conducted. Articles containing the keywords 'training' and 'assessment' are included and duplicates removed. The 87 included articles were downloaded as PDFs and organised in folders based on publication year and journal of origin.

The *Journal of Navigation* is an internationally peer-reviewed journal associated with Cambridge University Press. The journal publishes original papers on the science of navigation, covering every aspect of navigation, from the technical to the descriptive and historical. The first issue dates to 1948, which makes the journal a vast resource for understanding the science of navigation from a historical perspective. *WMU Journal of Maritime Affairs* is an international, peer-reviewed Springer journal associated with the World Maritime University in Malmö, Sweden. The journal covers subjects such as maritime safety, maritime energy, maritime administration, management and operations and marine environment protection. The journal gives special attention to human factors, the impacts of technology and policymaking, and it has a special section dedicated to Maritime Education and Training: The International Association of Maritime Universities (IAMU). The first issue dates to 2002. *TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation* is a peer-reviewed, open access journal associated published by Gdynia Maritime University, Gdynia, Poland. The journal was founded in 2007 and publishes original research and reviews contributing to the science of broadly defined navigation, aiming to promote young researchers and new fields of maritime sciences and technologies.

2.2 Approaching topics in MET

We adopted a mixed methods approach, which combines quantitative and qualitative approaches, to analyse these scientific publications. Figure 1 describes the pipeline for creating and analysing the corpus, both quantitatively and qualitatively. The corpus of 87 articles in the form of PDF files was imported to the text analysis module WordStat (v.9, Provalis Research). The texts were then mined using WordStat's topic
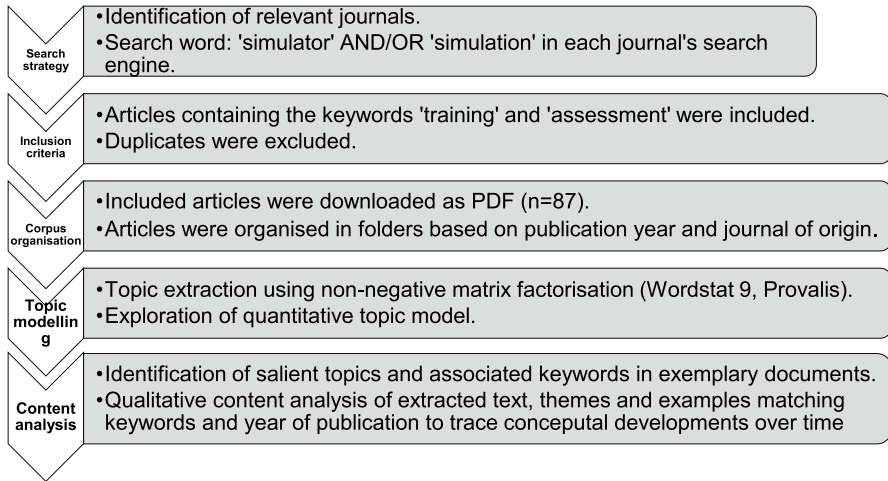
Fig. 1 A mixed methods approach to topic modelling of simulation-based education in MET

extraction feature, which combines natural language processing with statistical analysis to reveal its thematic characteristics.

First, we performed statistical topic modelling of the corpus to explore the most prominent topics related to training and assessment in simulation from our sample of journals for the period from 1961 to 2022. In this case, we created a model with 15 gross topics ($K = 15$). Although the choice of topic numbers (K) has been a subject in technical discussions, this decision ultimately depends on the researcher's interests (Jaworska & Nanda 2016, p. 383). Although modelling *too few* topics may produce themes that are too broad to be meaningful for a human reader, asking the model for *too many* topics can produce a highly fragmented and redundant model (see Ignatow & Mihalcea 2017, p. 209; Munksgaard & Enghoff 2018; Murakami et al. 2017).

Inspired by how scoping reviews have been used to approach conceptual developments in education (Major et al. 2018), we then conducted a more in-depth qualitative content analysis. Specifically, we analysed keywords from each topic retrieved from 'exemplary' segments of text to understand how the various topics appear in context over time, how themes were explicated and what kind of examples were used to represent various conceptual discussions by the authors. By exemplary texts, we here mean segments of text where the keywords for each topic figures prominently. In other words, the topic model *guided* our attention in the subsequent qualitative phase of the analysis. Here, our methods move from the language of machine learning and text mining of manifest content (so-called 'distant reading') to hermeneutics and questions about how the meanings of texts in simulation-based pedagogies in MET should be interpreted, e.g., 'close reading' (Gallagher 1992). In this phase, we leveraged our conceptual understanding of the field of simulation-based training in MET to investigate texts for patterns of meaning in more detail and meaningfully contextualize these (Russel and Ryan 2003; Major et al. 2018).

2.3 On topic modelling

Topic modelling is a probabilistic and heuristic method for building models of manifest content in large amounts of text. Whereas a manual approach for identifying topics may use coloured markers to highlight, code and group together the salient themes in a text (see Brett 2012), topic modelling uses a computer algorithm to sort through all the documents in a digitised corpus to create a higher-level abstract model of the structural relationships in these texts. The resulting automated clustering is based on word co-occurrences that are 'substantively meaningful' (Mohr & Bogdanov 2013). A commonly invoked metaphor is that topics are akin to 'bags of words' (Murakami et al. 2017). While a given topic is comprised by distributions of words, topic models are 'distributions of distributions (of words) across documents' (Shadrova 2021, p. 7) These patterns spatial proximity in text are rather easy to discover for computers, but harder to spot for human readers. In the language of machine learning, topic modelling is an unsupervised task: an inductive and data-driven approach to the objective properties of texts that does not require coding or annotation beyond the creation of the corpus itself. This approach can therefore avoid some of the biases that often characterise human readers (Underwood 2017, p. 19). Topic models are made from a probability distribution. As a computational method, topic modelling is guided by the assumption that meanings are relational and that words can be clustered in a meaningful way by the way they co-occur, 'regardless of syntax, narrative or location within a text' (Ignatow & Mihalcea 2017, p. 157). As such, these neither optimise nor exhaust interpretability but have the potential to meaningfully capture the 'aboutness' of texts (DiMaggio et al. 2013). Here, aboutness refers to what a given topic (collections of words and phrases that co-occur) is about, e.g., meaning in context. This is a task for human readers. As aptly described by Shadrova recently: 'topics are not direct derivatives of data, but interpretations of aspects of the data in context, where the context is provided by the scholar's knowledge of a subject and taxonomies and debates in the field' (2021: 4).

The modelling approach adopted in the current study combines natural language processing and a statistical method called factor analysis.[1] In WordStat, we pre-processed the data using its lemmatisation routine for English. This is an automated technique to process different forms of words into more canonical forms (e.g., plurals to singular, past tense verbs to present tense). Next, the segmentation level for the corpus of text was set at the level of paragraphs, reflecting the distribution of topics in typical scientific papers (a genre of documents that is quite long and where there is a need to compare the relative frequencies of topics). Using a random seed value, we computed a frequency matrix and extracted a smaller number of factors using a type of factor analysis called non-negative matrix factorisation (NNMF). For

---

[1] Other popular probabilistic methods include latent Dirichlet allocation (LDA) and latent semantic analysis (LSA). Ignatow, G., & Mihalcea, R. (2017). *An introduction to text mining: research design, data collection, and analysis*. Sage Publications. Peladeau and Davoodi (2018) offer a technical review of how factor analysis for topic extraction compares with LDA.

Table 1 Gross results of a topic model composed of 15 topics. Items are organised in the same rank order as the outputs from WordStat according to their factor number (NO, from the left). The TOPIC column contains the given label for each topic (WordStat suggests a name from keywords, here based on an algorithm; these have been edited for clarity). KEYWORDS displays the first 20 keywords for each topic. COHERENCE refers to the weighted average of correlated words associated with the topic. FREQ(UENCY) represents the total frequency of all the items that appear in the keyword's column. CASES displays the number of cases where at least one case for the keywords appears. Finally, % CASES gives the percentage of the same. *Being an artefact, Topic 9 is removed from the analysis below

NO	TOPIC	KEYWORDS	COHERENCE (NPMI)	FREQ	CASES	% CASES
1	MANAGING RESOURCES	RESOURCE; MANAGEMENT; LEADERSHIP; SKILLS; BRM; MARKERS; BEHAVIOURAL; TECHNICAL; TEAM; ERM; TECHNICAL SKILLS; RESOURCE MANAGEMENT; CREW RESOURCE MANAGEMENT; BEHAVIOURAL MARKERS; BRIDGE TEAM; BRIDGE RESOURCE MANAGEMENT; BRM TRAINING; ROOM RESOURCE MANAGEMENT	0.333	896	62	79.49%
2	ACADEMIC PERFORMANCE	IMPACTING; LICENCE; PRESENTEEISM; ACADEMIC; DIFFERENCE; REGIMENTAL; SCORE; PERCEIVED; AVERAGE; PPAP; PERFORMANCE; IDENTIFY; STUDENTS; ACADEMIC PERFORMANCE; LICENCE STUDENTS; PRESENTEEISM SCORE; AVERAGE PRESENTEEISM SCORE; IMPACTING THEIR ACADEMIC PERFORMANCE; LEVEL OF PRESENTEEISM; PERCEIVED ACADEMIC PERFORMANCE	0.491	1273	67	85.90%
3	RELATIVE MOTION	MOTION; RELATIVE; TRUE; MANOEUVRE; SPEED; VESSEL; PILOTS; PLOT; SWING; MANOEUVRES; TRACK; SUPPORT VESSEL; RELATIVE MOTION; RELATIVE PLOT; TRUE PLOT	0.246	824	65	83.33%
4	PORT CLUSTER	CLUSTER; FIRMS; CLUSTERS; REGIME; PORT; EDUCATION; TRAINING AND EDUCATION; PORT CLUSTER; LEADER FIRMS; PORT CLUSTERS; PORT RELATED; EDUCATION INFRASTRUCTU RE; TRAINING AND EDUCATION REGIME; QUALITY AND AVAILABILITY OF LABOUR; COLLECTIVE ACTION; IMPROVE THE QUALITY	0.127	492	58	74.36%

Table 1 (continued)

NO	TOPIC	KEYWORDS	COHERENCE (NPMI)	FREQ	CASES	% CASES
5	EYE TRACKING	TRACKING; GAZE; EYE; USER; SATISFACTION; SENTIMENT; USABILITY; EMOTIONAL; EYE TRACKING; GAZE TRACKING; EYE-TRACKING DATA; EYE GAZE; SPEECH RECORDING; PAPACHRISTOS ET AL; EYE-TRACKING SYSTEM; GAZE TRACKER; HEAD MOVEMENT; PERSONAL SATISFACTION; SENTIMENT OPINION	0.232	344	23	29.49%
6	HUMAN FACTORS	REVIEW; LITERATURE; HUMAN; FACTORS; MARITIME; CRM; LITERATURE REVIEW; HUMAN FACTORS; MARITIME EDUCATION; MARITIME INDUSTRY; CRM TRAINING; HUMAN ERROR; MARITIME ACCIDENTS; INTERNATIONAL MARITIME; MARITIME EDUCATION AND TRAINING	0.230	1026	68	87.18%
7	COMPUTER GAMES AND INSTRUCTIONAL DESIGN	GAME; GAMES; COMPUTER; INSTRUCTIONAL; EDUCATIONAL; ACCEPTANCE; COMPUTER GAMES; INSTRUCTIONAL DESIGN; GAME ATTRACTIVENESS; GAME DESIGN; AFFECTED BY THE PREVIOUS; AFFECTED BY THE LEVEL; GAME ACCEPTANCE AS EDUCATIONAL METHOD; GAME PLAY; SIMULATION GAMES; EDUCATIONAL GAMES; SIGNIFICANT LEVEL	0.217	525	45	57.69%
8	FIRE FIGHTING	FIRE; FIGHTING; SST; MADRAS; EVACUATION; ECR; PASSENGERS; FIRE FIGHTING; EVACUATION EQUIPMENT; EVACUATION EXERCISES; EVACUATION PROCESS; FIRE ALARM; FIRE FIGHTING TEAM; AFFECTED AREA; BREATHING APPARATUS; SAFETY EQUIPMENT	0.340	273	8	10.26%
9	METADATA*	CORE; CAMBRIDGE; HTTPS; ORG; TERMS; WWW; DOI; CORE TERMS; ORG CORE TERMS HTTPS	0.343	2704	69	88.46%
10	SIMULATOR TRAINING	SIMULATOR; INSTRUCTOR; BRIEFING; DEBRIEFING; EXERCISE; EXERCISES; SESSIONS; SESSION; TRAINEES; ROOM; SIMULATOR TRAINING; ENGINE ROOM; ENGINE ROOM SIMULATOR	0.323	1347	67	85.90%

Table 1 (continued)

NO	TOPIC	KEYWORDS	COHERENCE (NPMI)	FREQ	CASES	% CASES
11	POLLUTION VIOLATIONS	POLLUTION; VIOLATIONS; OIL.; ATTITUDE; FATIGUE; OIL POLLUTION; POLLUTION VIOLATIONS; POLLUTION PREVENTION; POLLUTION PREVENTION PRACTICES; OIL POLLUTION VIOLATIONS; POLLUTION INCIDENTS; NUMBER OF HOURS OF TRAINING; ATTITUDE TOWARDS POLLUTION PREVENTION PRACTICES; OIL POLLUTION PREVENTION AWARENESS; HOURS OF TRAINING; POLLUTION LAWS	0.341	342	25	32.05%
12	COLLISION AVOIDANCE, CERTIFICATION AND WATCHKEEPING	NAVIGATION; WATCHKEEPING; COLLISION; NAVIGATIONAL; STCW.; AVOIDANCE; SEAFARERS; COLLISION AVOIDANCE; CERTIFICATION AND WATCHKEEPING; CERTIFICATION AND WATCHKEEPING FOR SEAFARERS; STCW CODE; STANDARDS OF TRAINING; RADAR NAVIGATION; SAFE NAVIGATION; SAFETY OF NAVIGATION; STCW CONVENTION; WATCHKEEPING OFFICERS	0.198	635	70	89.74%
13	TRAINMOS II EUROPEAN COMMISSION	TRAINMOS; EUROPEAN; PROJECT; COMMISSION; MODULES; PROGRAMME; TRAINMOS II; EUROPEAN COMMISSION; TRAINMOS II PROJECT; EUROPEAN UNION; EDUCATIONAL PROGRAMME; TRAINING PROGRAMME; TRAINING MODULES; EUROPEAN COMMUNITIES; EUROPEAN PARLIAMENT; EUROPEAN TRANSPORT	0.302	249	32	41.03%
14	AUTHENTIC ASSESSMENT	AUTHENTIC; RELIABILITY; VALIDITY; ASSESSMENT; RUBRICS; ASSESSMENTS; EVIDENCE; CONSISTENCY; SEAFARER; COMPETENCE; AUTHENTIC ASSESSMENT; VALIDITY AND RELIABILITY; ASSESSMENT METHODS; AUTHENTIC ASSESSMENTS	0.413	1332	57	73.08%
15	THEMATIC BLOCK LEARNING	THEMATIC; BLOCK; TOPICS; CONTENT; PROGRAMME; THEMATIC BLOCK; LEARNING CONTENT; TEACHING LEARNING CONTENT; VISC PROGRAMME; THEMATIC BLOCK NATURE; THEMATIC BLOCKS; TOPICS INCLUDED; RESEARCH ACTIVITY; THEMATIC CONTENT; IMO PROGRAMME; VOCATIONAL SUBJECTS; OCCUPATIONAL STANDARD	0.353	216	18	23.08%

topic extraction, we included all words with a loading higher than a criterion of 0.30 (WordStat's default value).[2] Note that when modelling topics using factor analysis in WordStat, words may be associated with more than one factor, which 'more realistically represents the polysemic nature of some words as well as the multiple contexts of word usage' (Provalis Research 2022).

3 Results: a topic model of MET

Table 1 displays the solution for a model of the corpus, with a granularity of 15 topics ($K=15$). Of these 15 topics, we consider 14 to be 'substantively' meaningful (Murakami et al. 2017, p. 250), based on our common knowledge about the domain in question. Notably, topic nine is clearly 'chimeric' (Ignatow & Mihalcea 2017, p. 212), an artefact based on keywords from metadata about scientific articles that is automatically generated when extracting documents from databases. Therefore, we discarded this topic from our qualitative investigation. A further review of the remaining models, considering our research questions, narrowed the scope of our hermeneutic inquiry to four topics. These are discussed in detail below, along with a brief justification for why the other nine topics were discarded from a thorough qualitative inquiry.

Again, we would like to emphasise that a key criterion for assessing a given topic model is whether it reveals something interesting about texts, given a particular set of research questions. When assessing a topic model beyond the 'aboutness' of the corpus, topic modellers frequently discard topics that are orthogonal to their research interests or even merge overlapping ones. Here, we were primarily interested in the questions of the following: (1) What are the most prominent topics related to training and assessment in simulation in a corpus of articles from selected maritime journals from 1961 until 2021, and (2) What kind of central themes are used to describe simulator-based training and assessment in the corpus, and how do these themes and their examples develop over time?

Topic models can be supported by identifying compelling and meaningful patterns in large amounts of textual data in ways that challenge the researchers' intuitions from the onset of a critical inquiry. However, in contrast to a conventional approach to content analysis, in which the reader closely peruses documents to gain an understanding of their subject matter, the text-mining approach we adopt here makes an inductive catalogue of the content first, displacing interpretative work to *after* the computational procedure has been performed. This inductive step *reduces* the risk that researchers will project their own assumptions about the field on the materials. The details about this process are described in the next section.

---

[2] Increasing the cutoff-value reduces the number of words in the column KEYWORDS in Table 1, keeping only words that are more representative for each topic (and vice versa). Other defaults include removing phrases ending with words on WordStat's default exclusion list for English and limiting phrases to a minimum of two and a maximum of five words.

4 Qualitative exploration

In this section, we offer a qualitative exploration of our topic model of simulation-based MET. Because our topic model is an inductive approach to the objective properties of the texts in the corpus, that is, their 'aboutness' (DiMaggio et al. 2013), we seek to discover meaningful relationships between words that are clustered together and co-occur within the specific topics. As mentioned, the task of common-sense understanding of words in context is something machines are not well equipped to do for a variety of reasons, and topic models therefore require external validation by a human reader (Shadrova 2021). Central to addressing our two research questions, we will analyse the following four topics: *Topic 1—managing resources*, *2—academic performance*, *10—simulator training* and *14—authentic assessment.* Note that in this qualitative review, we order our discussion of these four topics according to which one *has the highest percentage of cases with at least one of the items listed in the keyword's column* (rightmost column, Table 1). This rank order lets us better address our second research question, that is, what kind of central themes are used to describe simulator-based training and assessment in the corpus, and how these themes develop over time. In other words, we use the percentage of cases as a proxy for ordering topics in the corpus from the most extensive discursive tradition in the corpus to the least.

After reviewing all the topics in our model, topics 3–5, 7–8, 11–14 and 15, emerged as orthogonal for our research questions. The justification for not including these is as follows:

- Topics 3, 4, 8, 11 and 12 revolve around prominent *content* in maritime training and education, such as relative motion, navigation in ports, firefighting, pollution and collision avoidance. These subjects do not meaningfully pertain to how training and assessment are conceptualised through the use of simulation-based training.
- The keywords for *topic 5—Eye tracking* reveal links between eye-tracking technologies, user, usability, satisfaction and sentiment opinion. Hence, the topic relates to usability testing rather than to matters of learning in simulator environments.
- *Topic 6—Human factors* is a prominent topic in the data and appears in 68 cases/articles. The topic stands out among the others because it denotes a field of research rather than an aspect of simulation. A qualitative content analysis of keywords shows that a large portion of the cases concern bibliographic references, revealing the strong position of human factors approaches in MET.
- We also left out *Topic 7—Computer games and instructional design* because the keywords in this topic were represented by only one scientific article and, hence, was too specific to be considered a theme in the broader research field of MET.
- The justification for leaving out *Topic 13—TRAINMOS II European Commission* is because these keywords appear in one article outlining TRAINMOS II,

Table 2 Topic model results from Topic 1—managing resources, documenting each case, text, match and file with article location

	A	B	C	D	E	F	G	H
1	Case #	Text	Matching	Case	FILE	LOCATION	CREATED	Variable
2	72	The following terms were used in the search: non-technical skills, human factors, bridge resource management /BRM, bridge team management/BTM, engine room resource management (ERM), crew resource management/CRM, simulator-based training and maritime. The word maritime was included in all the searches in combination with at least one of the other terms. Different databases were used to identify relevant research in the field of interest. An initial search was performed using the ORIA database available to researchers at the Norwegian University of Science and	MANAGEMENT; TEAM; RESOURCE; ERM; TECHNICAL; BRM; SKILLS	Case #72	2018JOMA1	2018	DOCUMENT	105
3	72	), now requiring all ship officers to undergo a leadership and teamwork training and demonstrate knowledge of bridge and engine room resource management principles (BRM and ERM respectively) to be certified or to renew their certificates. This has led to an increased interest in non-technical skills in the maritime industry in general and among researchers. Still, there is no updated overview of studies in the field. This article will contribute in this respect, as it aims to review research from 2010 to 2017 that identifies applicable factors relevant to the non-technical training of maritime officers.	TECHNICAL; SKILLS; MANAGEMENT; RESOURCE; ERM; BRM; LEADERSHIP	Case #72	2018JOMA1	2018	DOCUMENT	25

the continuation of a large project on transport and sustainability funded by the European Commission on Supporting Motorways of the Sea (MOS).

- *Topic 15—Thematic block learning* is also absent from our qualitative review. Keywords comprising this particular topic refer to a specific pedagogical model in curriculum design through a case study from the Latvian Maritime Academy (Kalnina & Priednieks 2017). The research article itself does not address simulation-based learning directly and is therefore not subject to a qualitative review.

After deciding on the four topics based on the highest percentage of cases: managing resources, academic performance, simulator training and authentic assessment, we investigated the statistical text mining results in detail (see Table 2). For each case-number (for example #72, leftmost column), we analysed the extracted 'text' for each time the case occurred. By analysing the themes and their examples in the text-excerpts, matching keywords and year of publication for each case, we explored how the themes and their examples developed over time.

In the analysis that follows, keywords in ALL CAPS are used as notations for representing the 20 chosen and specific keywords from each of the four topics in the model (Table 1), as well as the examples and citations which are applied from the most relevant 'cases' and their 'text-extractions' (Table 2).

4.1 Topic 1—managing resources

1	Managing resources	RESOURCE; MANAGEMENT; LEADERSHIP; SKILLS; BRM; MARKERS; BEHAVIOURAL; TECHNICAL; TEAM; ERM; TECHNICAL SKILLS; RESOURCE MANAGEMENT; CREW RESOURCE MANAGEMENT; BEHAVIOURAL MARKERS; BRIDGE TEAM; BRIDGE RESOURCE MANAGEMENT; BRM TRAINING; ROOM RESOURCE MANAGEMENT

The keywords for *Topic 1—Managing resources* has a focus on *categorising skills*. The keywords comprising this topic combine descriptions of various domains (CREW RESEROUCE MANAGEMENT, BRIDGE RESOURCE

MANAGEMENT/BRM, ENGINE ROOM MANAGEMENT/ERM), actors involved (LEADERSHIP, BRIDGE TEAM) and properties of technical skills (BEHAVIOURAL, TECHNICAL, MARKERS). The topic is prominent in the corpus and can be found across 62 cases/articles, with keywords that make up the topic being mentioned 896 times across all three journals for the studied period. A qualitative content analysis analysing how these keywords appear in context ($n = 896$) reveals a corpus of studies on a specific type of team training, labelled crew resource management (CRM), bridge resource management (BRM), engine resource management (ERM) and team training, which appears in combination with what is commonly known as 'nontechnical skills', that is, skills that are considered either cognitive or communicative. Although CRM originated from aviation psychology in the late 1060 s, Praetorius et al. (2020) describes how the concept was transferred to the maritime domain in the early 1990s:

> A first maritime version of the CRM course package focused on bridge operations was developed 1992 (Hayward, Lowe, and Thomas, 2019). In 1995, the IMO introduced the concept of Bridge Resource Management (BRM) - the effective use and allocation of all resources available on the bridge - into the STCW Code (Chauvin et al. 2013). Since then, the concept has been transferred to other departments onboard as the development of courses for ERM and MRM has continued. Through the latest amendments to the Code, NTS knowledge has become mandatory (IMO, 2017).

Because BRM and MRM training are mandatory parts of the current MET system, the frequent occurrence of cases in our corpus of articles shows a research field dedicated to determining if current training and assessment practices in BRM and MRM courses fulfil their purpose of ensuring high safety standards. Notably, we find the first occurrence in our corpus more than a decade before the first BRM course was developed. The first article explicitly mentioning technical skills was published in 1980 in the *Journal of Navigation*. In this article by Hammell and Gardenier, a deck officer behavioural database is assembled 'to define the ship handling skills to be achieved by masters and chief mates, in the form of Specific Functional Objectives (SFO). The SFOs represent detailed goals of the training system. This developmental effort concentrated on the master's position, pertaining to manoeuvring of the vessel for several different classes of vessels. Although there are few studies addressing this topic between 1981 and 2013 across the three different journals, a new wave of articles on technical and nontechnical skills emerges between 2015 and 2019, mainly in JOMA. In this period, we find broad discussions about the limitations of measuring technical skills (Wild 2011; Bhardwaj 2013; Felsenstein et al. 2013). The following studies emphasise 'soft skills', such as situation awareness and communication, as important for collaborative navigation practices in bridge teams (Wu et al. 2015; Röttger et al. 2016; Saeed et al. 2017; Kandemir et al. 2018).

4.2 Topic 2—academic performance

2	Academic performance	*IMPACTING; LICENCE; PRESENTEEISM; ACADEMIC; DIFFERENCE; REGIMENTAL; SCORE; PERCEIVED; AVERAGE; PPAP; PERFORMANCE; IDENTIFY; STUDENTS; ACADEMIC PERFORMANCE; LICENCE STUDENTS; PRESENTEEISM SCORE; AVERAGE PRESENTEEISM SCORE; IMPACTING THEIR ACADEMIC PERFORMANCE; LEVEL OF PRESENTEEISM; PERCEIVED ACADEMIC PERFORMANCE*

The keywords for *Topic 2—academic performance* revolve around measurements, including terms like IMPACTING; SCORE; AVERAGE; PERFORMANCE; PRESENTEEISM SCORE; AVERAGE PRESENTEEISM SCORE; LEVEL OF PRESENTEEISM and PERCEIVED ACADEMIC PERFORMANCE, that is, different kinds of metrics for performance evaluations. Pursuing a qualitative content analysis and analysing each case occurrence associated with the topic ($n = 1273$), the topic emerges as less instrumental. Rather, these cases focus on exploring how to licence students based on performance tests in simulator environments. Furthermore, *academic performance* emerges as one of the most prominent topics in our model and can be found across 67 articles, spanning publications in *Journal of Navigation* in the early 1960s towards contemporary articles in *TransNav* and *WMU JOMA*. In the first wave of articles from 1961 to 1966 (Quick 1961; Burger and Corbet 1963; The Training of Navigators 1966), simulator-based learning is discussed in relation to technical issues concerning radar equipment and work practice at sea. Early work highlights the importance of students gaining practical experience to proficiently carry out work at sea (Quick 1961), but also the need for nautical instructors to keep their practical experiences of seamanship, as well as their technological proficiencies, up to date and relevant to adequately teach the next generation of seafarers (The Training of Navigators 1966). In these works, which are disseminated by instructors working with the early development of simulator-based training, the learning process is cast in terms of common-sense 'folk pedagogy' (Bruner 2020): learning is considered to happen automatically as trainees get exposed to professional practices. An illustrative example of such discourse is found in the following phrase from 1961:

> The teaching and examination have necessarily to be somewhat academic. It is assumed that students, after satisfactorily completing the course and being appointed to a radar-fitted vessel, will gain practical experience at sea, particularly practice in plotting target echoes in clear weather, so that by the time they are called on to deal with conditions of poor visibility, they will be sufficiently adept at the techniques to be able to carry out satisfactorily the routines required. (Quick 1961)

It is noteworthy, however, that Quick (1961) does not explicitly mention learning when discussing the trajectory from being a student completing a course towards being a seafarer with enough practical experience to handle challenging conditions

at sea. Although the topic appears in few studies between 1966 and 2004, a new wave of articles on academic performance and the licensing of students emerges early in the new millennium. Without exception, these articles are published in *WMU JOMA* (founded in 2002) and *TransNav* (founded in 2007). Ircha and Balsom (2005) discuss simulator-based learning together with the new educational technologies entering the MET system at the time. This includes e-mails, web-based materials, automated quizzes and new forms of organising professional training through online lectures and web-based courses, alongside advanced high-fidelity simulators (Ircha and Balsom 2005; Benedict et al. 2006; Zažeckis et al. 2009). Here, emphasis is on pedagogical models in simulator-based learning that centre on the student rather than instructor (Ircha and Balsom 2005). After 2012, studies focusing on measuring student satisfaction in relation to academic performance appear (Vasilakis and Naikitados 2012; Castells et al. 2016). We can also see a new focus on the validity and reliability of performance tests in simulators used for licensing students in MET (Ghosh et al. 2016; Balaji and Venkadasalam 2017; Conceição et al. 2017; Øvergård et al. 2017). Conceição et al. (2017), for instance, suggest an approach to assessment influenced by human factors research in aviation, here by developing an assessment model for identifying and rating naval cadets' nontechnical skills. Ghosh et al. (2016), on the other hand, advocate for so-called 'authentic assessment', that is, a holistic approach with a focus on realistic, work-related tasks and test contexts. Automated performance assessment is another approach suggested for improving validity and reliability when issuing certificates in a simulator environment (Øvergård et al. 2017).

4.3 Topic 10: simulator training

10	*Simulator training*	SIMULATOR; INSTRUCTOR; BRIEFING; DEBRIEFING; EXERCISE; EXERCISES; SESSIONS; SESSION; TRAINEES; ROOM; SIMULATOR TRAINING; ENGINE ROOM; ENGINE ROOM SIMULATOR

The keywords for *Topic 10—Simulator training* demarcate simulation training as a distinct activity. The keywords comprising the topic pick out various parts of simulation events, here in terms of their spatial (SIMULATOR; ROOM; ENGINE ROOM) and temporal properties (e.g., SESSIONS, BRIEFING; TRAINING EXERCISE; DEBRIEFING), and some of the main actors involved in the educational practice (INSTRUCTOR; TRAINEES). Notably, the topic is found across 67 cases (articles) representing the entire timespan in the corpus from 1961 to 2018. Except for those keywords in Topic 9, which is an artefact from the corpus' metadata, the keywords from this topic also appear the most frequently in the corpus (1,347 times). A central theme are descriptions of simulation training as a highly structured learning event. An illustrative quote can be found in an article from a 2006 article in the *Journal of Maritime Affairs*, containing the terms SESSION; INSTRUCTOR; EXERCISE; and SIMULATOR:

One of the most important parts of the simulator exercise is the evaluation of the students' results by the instructor both during and after the training session. This should be performed in two ways: first, during the exercise run to ensure that the training objective can be achieved and second after exercise completion in order to give the students an indication of their performance during the simulation run. (Benedict et al. p. 16)

Clearly, this topic provides an indispensable descriptive frame for the authors of these articles on MET. It makes it possible for researchers to describe and represent simulation-based pedagogies as a coherent and durable set of pedagogic relationships and events that can be apt subjects for scientific discourse. As such, the topic plays a role in statements that explicate and highlight precisely *how* the training event in question is organised, which help explain its many appearances across most of the articles in our corpus.

4.4 Topic 14—authentic assessment

14	Authentic assessment	AUTHENTIC; RELIABILITY; VALIDITY; ASSESSMENT; RUBRICS; ASSESSMENTS; EVIDENCE; CONSISTENCY; SEAFARER; COMPETENCE; AUTHENTIC ASSESSMENT; VALIDITY AND RELIABILITY; ASSESSMENT METHODS; AUTHENTIC ASSESSMENTS

[…] good assessment (formative or summative) has to be valid, reliable, practical, developmental, manageable, cost-effective, fit for purpose, relevant, authentic, closely linked to learning outcomes and fair. (Zažeckis et al. 2009)

The keywords for *Topic 14—Authentic assessment* focus on scientific values such as VALID, ASSESSMENT, REALIABLE, PRACTICAL, DEVELOPMETAL, MANAGEEABLE, COST-EFFECTIVE, FIT FOR PURPOSE, RELEVANT and AUTHENTIC. A qualitative content analysis analysing each data co-occurrence associated with the topic ($n = 1,332$) displays a corpus of studies that are critical for current assessment practices in MET and that are concerned with developing and validating assessment methods towards valid, reliable and consistent tests of competence in simulator environments. Like the topic *Academic Performance*, the topic *Authentic Assessment* figures prominently and is found in 57 articles, spanning from the first article published in *Journal of Navigation* in 1962 to an article published in *WMU JOMA* in 2018. Although the first articles in the 1960s mostly address assessments in the simulator as a rather straightforward matter (Weekes 1962; The Training of Navigators 1966), the complexities of simulator-based assessments become thematically central throughout the 1980s (Taylor 1998). At this time, reflections about the authenticity of test situations during individual driving tests are related to the everyday collaborative navigation practices of bridge teams. The studies that follow emphasise validity, reliability and fairness in relation to simulator-based assessments (Zažeckis et al. 2009; Malik and Zafar 2015; Orlandi et al. 2015; Ghosh et al.

2016, 2017; Sellberg 2017; Ernstsen and Nazir 2018). Although some of the studies are practitioner driven and practice oriented, investigating simulator-based assessment in their own MET institutions (Zažeckis et al. 2009; Malik and Zafar 2015), others are more theory driven and empirically grounded (Orlandi et al. 2015; Ghosh et al. 2016, 2017; Sellberg 2017; Ernstsen and Nazir 2018). Nevertheless, in these studies, there is consensus that simulator-based assessments in MET need to be improved if they are to be perceived as legitimate tests of professional competence.

5 Discussion and conclusion

Methodologically, the current paper contributes to research on simulation-based training in higher education for the professions by adopting a novel approach, combining quantitative topic modelling and qualitative content analysis of the most central themes in a corpus of premier scholarship on MET. The first review of its type in this area, the approach makes visible implicit conceptual notions about simulation-based education and how these notions have developed and changed over time. In other words, the quantitative topic model, combined with in depth, 'close reading', represents an inductive and data-driven approach to the 83 texts comprising our corpus. Consequently, a text mining method is particularly suitable for literature reviews because it allows us to automate certain aspects of process, revealing features of discourse that a human readers might miss (Fabbri et al. 2013; Thomas et al. 2011; Audrin and Audrin 2022). In our case, the topic model became an interpretative resource for further inquiry on discourse on simulator-based training and assessment. In line with the reasoning of DiMaggio et al. (2013), the probability distribution in our topic model was qualitatively validated to meaningfully capture the 'aboutness' of the prominent texts that conceptualise simulator-based learning in MET. While undoubtably reductionistic in itself, our text-mining approach provided an inductive catalogue of content first, displacing interpretative work to the aftermath of the computational procedure. So while the topic model is created inductively, this next interpretative step is actually an abductive process, where contextualization through the scholar's own frame of reference play an important role (Shadrova 2021, p. 19).

The strength of augmenting close reading with quantitative topic modeling, is that it does not primarily rely on the researcher's own predefined concepts and categories. Being inductive and data-driven, the topic model grounded our critical analysis of discursive trends in the field. It is not a given that this text-mining approach would result in the same themes that a more conventional in-depth reading would yield. Neither is the model to be taken at face-value, as our qualitative corroboration of the model did not find each of the 15 topics equally informative. While the search words initially chosen certainly constrains scope of the corpus, the quantitative model should not be considered as merely an artefact of our search-procedure. Rather it should be considered as an epistemic resource that adds to our understanding of how salient topics in the MET-field has developed over time, extending the

capacities of the human reader to identify and zoom in on fine-grained particulars across texts.

Our results indicate that the most prominent topics related to training and assessment in simulation from the selected maritime journals *TransNav*, *WMU*, *JOMA* and *Journal of Navigation* from 1961 until today are managing resources, academic performance, simulator training and authentic assessment. Notably, the 'aboutness' of these prominent topics reveals how the conceptualisation of simulator-based training and assessment is characterized by an instrumental approach, suggesting that the concept of learning is underdeveloped in the literature. Moreover, while the text-mining approach revealed that learning as a topic in maritime simulator-based training and assessment first occurs in the literature around 1980, it does not emerge as a central and explicit theme before 2012. Thus, the implications and insights garnered from central maritime journals for building up an academic field of simulator-based training display a focus on performance, authenticity in the assessment methods, management of resources and the simulator as a setting for training. Consequently, investigating what prominent topics related to training and assessment might be influential recommendations for developing future research on topics related to the scientific field of education and towards investigating the concept of learning in the context of MET.

The topic model also helps us address our second research question: to qualitatively examine what kinds of central themes are used to describe simulator-based training and assessment in the corpus and how these themes and their examples have developed over time. Through detailed content analysis, our findings shows that historically, the concept of learning is not explicitly fleshed out in academic publications on MET between 1961 and 2012. Instead, the field has relied on what Bruner (1996) describes as 'folk pedagogy', a set of lay theories and implicit assumptions about how students learn, without a clear theoretical foundation in the learning sciences and overt concerns about didactic competence. As Sellberg (2017) documents, MET is a research field historically dominated by human factors research. In our quantitative analysis, we can see how these studies are focusing on safety-related themes also when studying training and assessment in BRM and MRM courses (e.g., Felsenstein et al. 2013; Praetorius et al. 2020). Hence, our content analysis reveals that the central themes to describe simulator-based training and assessment are more oriented towards *learning objectives* as a product, than on the *learning activity* as a process. Gaining momentum over the last decade, a small but growing corpus of studies now also aims to scientifically investigate learning *activities* empirically, for instance, by examining simulator-based training and assessment practices (Benedict et al. 2006; Sellberg and Wiig 2020; Sellberg et al. 2022). Still, few studies on maritime training and assessment explicitly adopt process-based theories of learning. However, MET should be enriched by research that can advance the field in terms of what domains to train (and how to train them), ensuring that the training and assessment activities that take place in simulator environments fulfil those learning objectives connected to high-risk industries.

Moreover, our text mining approach suggests that reflexive conceptualisations of learning in maritime simulations appears to be mainly connected to the *WMU Journal of Maritime Affairs* and its IAMU section, a part of the journal dedicated to publishing research from the International Association of Maritime Universities. Because our investigation was based on data from three international, peer-reviewed

journals within the maritime field, this limitation to domain-specific journals might explain why the concept of learning makes such sparse appearances in the earliest research on simulator-based learning in MET. Certainly, valuable conceptualisations of simulation-based learning might be found in other journals, edited books and publication venues that are not peer reviewed, such as grey literature and content on social media. Still, the chosen journals are historically recognised as core publications within the scholarship on maritime training and assessment, thus providing an interesting repository for intellectual history about how simulator-based learning has been conceptualised by researchers. Moreover, while compiling the corpus in the data-collection phase, we chose to highlight 'training' and 'assessment' as the search words, rather than explicitly searching for the concept of 'learning'. This was based on insights from previous reviews by Sellberg (2017) and Wahl and Kongsvik (2018) that have documented how the sociology of knowledge in the field of MET emphasises training and assessment for international certification schemes. Here, our intention was to gather this assortment of studies and explore how learning is framed historically in peer-reviewed publications on MET. The mixed-methods approach to our corpus from three notable journals helps us better understand the historical transformation of MET from a vocational programme towards an academic education. These findings have the potential to inform the educational preparation of future professionals within other safety–critical domains and to raise our awareness about the epistemic presuppositions that inform simulation-based learning activities as tools for professional learning in higher education.

Declarations

Competing interests The authors declare no competing interests.

References

Audrin C, Audrin B (2022) Key factors in digital literacy in learning and education: a systematic literature review using text mining. Educ Inf Technol 27:7395–7419

Balaji R, Venkadasalam S (2017) Developing a marine engineering centre of excellence for competency-based training. WMU J Marit Aff 16(2):293–311

Bhardwaj S (2013) Technology, and the up-skilling or deskilling conundrum. WMU J Marit Aff 12(2):245–253

Benedict K, Baldauf M, Felenstein C, Kirchhoff M (2006) Computer-based support for the evaluation of ship handling exercise results. WMU J Marit Aff 5(1):17–35

Brett MR (2012) Topic modelling: a basic introduction. J Digit Humanit 2(1):1–2

Bruner J (2020) 2.Folk Pedagogy. In *The culture of education* (44–65). Harvard University Press

Burger W, Corbet A (1963) Radar simulator courses. J Navig 16(2):253–255. https://doi.org/10.1017/S0373463300019706

Castells M, Ordás S, Barahona C, Moncunill J, Muyskens C, Hofman W, Skorokhodov S (2016) Model course to revalidate deck officers' competences using simulators. WMU J Marit Aff 15(1):163–185

Chauvin C, Lardjane S, Morel G, Clostermann J-P, Langard B (2013) Human and organisational factors in maritime accidents: analysis of collisions at sea using the HFACS. Accid Anal Prev 59(0):26–37. https://doi.org/10.1016/j.aap.2013.05.006

Conceição VD, Basso JC, Lopes CF, Dahlman J (2017) Development of a behavioural marker system for rating cadet's non-technical skills. TransNav:Int J Mar Navig Safety Sea Transp 11(2):69

DiMaggio P, Nag M, Blei D (2013) Exploiting affinities between topic modelling and the sociological perspective on culture: application to newspaper coverage of US government arts funding. Poetics 41(6):570–606. https://doi.org/10.1016/j.poetic.2013.08.004

Ernstsen J, Nazir S (2018) Consistency in the development of performance assessment methods in the maritime domain. WMU J Marit Aff 17(1):71–90

Fabbri S, Hernandes E, Di Thommazo A, Belgamo A, Zamboni A, Silva C (2013) Using information visualization and text mining to facilitate the conduction of systematic literature reviews. In: Cordeiro J, Maciaszek LA, Filipe J (eds) Enterprise information systems. Lecture notes in business information processing, vol 141. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-40654-6_15

Ferreira-Mello R, André M, Pinheiro A, Costa E, Romero C (2019) Text mining in education. WIREs Data Mining Knowl Discov 9(6):e1332. https://doi.org/10.1002/widm.1332

Felsenstein C, Benedict K, Baldauf M (2013) Maritime safety and security challenges–3D simulation based training. TransNav: Int J Mar Navig Safety Sea Transp 7(3):327–336

Gallagher S (1992) Hermeneutics and Education. SUNY Press

Ghosh S, Bowles M, Ranmuthugala D, Brooks B (2016) Authentic assessment in seafarer education: Using literature review to investigate its validity and reliability through rubrics. WMU J Marit Aff 15(2):317–336

Ghosh S, Bowles M, Ranmuthugala D, Brooks B (2017) Improving the validity and reliability of authentic assessment in seafarer education and training: a conceptual and practical framework to enhance resulting assessment outcomes. WMU J Marit Aff 16(3):455–472

Ignatow G, Mihalcea R (2017) An introduction to text mining: research design, data collection, and analysis. Sage Publications, Thousand Oaks

Ircha MC, Balsom MG (2005) Educational technology: enhancing port training. WMU J Marit Aff 4(2):211–225

Jaworska S, Nanda A (2016) Doing well by talking good: a topic modelling-assisted discourse study of corporate social responsibility. Appl Linguis 39(3):373–399. https://doi.org/10.1093/applin/amw014

Kandemir C, Soner O, Celik M (2018) Proposing a practical training assessment technique to adopt simulators into marine engineering education. WMU J Marit Aff 17(1):1–15

Kalnina R, Priedieks V (2017) Proficiency improvement method in maritime education. WMU J Marit Aff 16(1):139–159

Lavissière A, Sohier R, Lavissière MC (2020) Transportation systems in the Arctic: a systematic literature review using textometry. Transp Res A Policy Pract 141:130–146. https://doi.org/10.1016/j.tra.2020.09.00310.1016/j.tra.2020.09.003

Major L, Warwick P, Rasmussen I, Ludvigsen, Cook V (2018) Classroom dialogue and digital technologies: a scoping review. Educ Inf Technol 23:1995–2028. https://doi.org/10.1007/s10639-018-9701-y

Malik A, Zafar N (2015) Applications of simulation technology-pitfalls and challenges. TransNav:Int J Mar Navig Safety Sea Transp 9(3):391

Manuel ME (2017) Vocational and academic approaches to maritime education and training (MET): Trends, challenges, and opportunities. WMU J Marit Aff 16(3):473–483

Mohr JW, Bogdanov P (2013) Introduction—topic models: what they are and why they matter. Poetics 41(6):545–569. https://doi.org/10.1016/j.poetic.2013.10.001

Munksgaard R, Enghoff O (2018) Køn og metodevalg blandt samfundsvidenskabelige specialeskrivende. Dansk Sociologi 29(1):63–95

Munn Z, Peters MD, Stern C, Tufanaru C, McArthur A, Aromataris E (2018) Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 18(1):1–7

Murakami A, Thompson P, Hunston S, Vajn D (2017) 'What is this corpus about?': Using topic modelling to explore a specialised corpus. Corpora 12(2):243–277. https://doi.org/10.3366/cor.2017.0118

Orlandi L, Brooks B, Bowles M (2015) A comparison of marine pilots' planning and manoeuvring skills: uncovering mental models to assess shiphandling and explore expertise. J Navig 68(5):897–914

Øvergård KI, Nazir S, Solberg AS (2017) Towards automated performance assessment for maritime navigation. TransNav, Int J Mar Navig Safety Sea Transp 11(2):43

Peladeau N, Davoodi E (2018) Comparison of latent Dirichlet modelling and factor analysis for topic extraction: a lesson of history. HICSS

Provalis Research (2022) WordStat v9.0: content analysis and text mining – manual. Provalis Res

Praetorius G, Hult C, Österman C (2020) Maritime resource management: current training approaches and potential improvements. TransNav, Int J Mar Navig Safety Sea Transp 14(3):573–584

Quick JH (1961) The radar simulator and collision avoidance at sea. J Navig 14(1):64–68

Russel BH, Ryan GW (2003) Techniques to identify themes in qualitative data. Field Methods 15(1):85–109

Röttger S, Vetter S, Kowalski JT (2016) Effects of a classroom-based bridge resource management training on knowledge, attitudes, behaviour, and performance of junior naval officers. WMU J Marit Aff 15(1):143–162

Saeed F, Wall A, Roberts C, Riahi R, Bury A (2017) A proposed quantitative methodology for the evaluation of the effectiveness of Human Element, Leadership and Management (HELM) training in the UK. WMU J Marit Aff 16(1):115–138

Sellberg C (2017) Simulators in bridge operations training and assessment: a systematic review and qualitative synthesis. WMU J Marit Aff 16(2):247–263

Sellberg C, Wiig AC (2020) Telling stories from the sea: facilitating professional learning in maritime post-simulation debriefings. Vocat Learn 13(3):527–550

Sellberg C, Wiig AC, Säljö R (2022) Mastering the artful practice of navigation: the situated endorsement of professional competence in post-simulation evaluations. Stud Educ Eval 72:101111

Shadrova A (2021) Topic models do not model topics: epistemological remarks and steps towards best practices. J Data Mining Digit Human. https://doi.org/10.46298/jdmdh.7595

Taylor DH (1998) Rules and regulations in maritime collision avoidance: new directions for bridge team training. J Navig 51(1):67–72

The Training of Navigators (1966) Discussion. J Navig 19(3):371–391. https://doi.org/10.1017/S0373463300047470

Thomas J, McNaught J, Ananiadou S (2011) Applications of text mining within systematic reviews. Res Syn Meth 2(1):1–14. https://doi.org/10.1002/jrsm.27

Thompson P, Batista-Navarro RT, Kontonatsios G, Carter J, Toon E, McNaught J et al (2016) Text mining the history of medicine. PLoS One 11(1):e0144717. https://doi.org/10.1371/journalpone.0144717

Underwood T (2017) A genealogy of distant reading. DHQ: Digit Human Q 11(2)

Vasilakis P, Naikitados N (2012) A methodological framework for evaluating maritime simulation. TransNav, Int J Mar Navig Safety Sea Transp 6(11):27–32

Wahl AM, Kongsvik T (2018) Crew resource management training in the maritime industry: a literature review. WMU J Marit Aff 17(3):377–396

Weekes MWM (1962) I—The use of simulators for radar training. J Navig 15(3):262–267

Wiig AC (2023) Hvordan utforske et ukjent forskningsfelt? En systematisk forskningsoversikt for artikkel. In TS Prøitz (eds) Forskningsoversikter i utdanningsvitenskap. Systematikkog kreativitet. (Translated from Norwegian: How to explore an unknown research field? A systematic research review for article. In TS Prøitz (eds) Research overviews in educational science. Systematics and creativity). Fagbokforlaget 105–128

Wild CRJ (2011) The paradigm and the paradox of perfect pilotage. J Navig 64(1):183–191

Wu Y, Miwa T, Shimamoto K et al (2015) Development of quantitative team performance evaluation method for ERM. WMU J Marit Aff 14:333–347. https://doi.org/10.1007/s13437-015-0088-y

Zažeckis R, Bartusevičiene I, Maksimavicius R, Sencila V, Rupšiene L, Kalvaitiene G, Baumler R (2009) New capabilities of the NTPRO 4000 full mission ship handling simulator in the assessment and evaluation processes at Lithuanian Maritime Academy. TransNav, Int J Mar Navig Safety Sea Transp 89–94